W.K. Kellogg Foundation
Logic Model Development Guide

Using Logic Models to Bring Together Planning, Evaluation, and Action

Logic Model Development Guide

To help people help themselves through the practical application of knowledge and resources to improve their quality of life and that of future generations.

Updated January 2004

W.K. Kellogg Foundation
One East Michigan Avenue East
Battle Creek, Michigan 49017-4012
www.wkkf.org

Contents

Chapter 4

Introduction

If you don't know where you're going, how are you gonna' know when you get there?
–Yogi Berra

In line with its core mission – *To help people help themselves through the practical application of knowledge and resources to improve their quality of life and that of future generations* – the W.K. Kellogg Foundation has made program evaluation a priority. As our staff and grantees work on a spectrum of social improvement programs, the need for shaping and contributing to the body of knowledge regarding evaluation becomes increasingly clear. Our first guide, the *W.K. Kellogg Foundation Evaluation Handbook,* was published in 1998, and has been made available to nearly 7,500 people. The *Evaluation Handbook* is a practical, step-by-step manual for conducting evaluations. With the *Handbook,* we introduced the concept of the **program logic model** and the ways in which applying this concept has added value to our own work.

The program logic model is defined as a picture of how your organization does its work – the theory and assumptions underlying the program. A program logic model links outcomes (both short- and long-term) with program activities/processes and the theoretical assumptions/principles of the program.

The *W.K. Kellogg Foundation Logic Model Development Guide,* a companion publication to the *Evaluation Handbook,* focuses on the development and use of the program logic model. We have found the logic model and its processes facilitate thinking, planning, and communications about program objectives and actual accomplishments. Through this guide, we hope to provide an orientation to the underlying principles and language of the program logic model so it can be effectively used in program planning, implementation, and dissemination of results.

The premise behind this guide – and our view of the role of evaluation in programming – is simple: Good evaluation reflects clear thinking and responsible program management. Over the years, our experience in using logic models in initiatives such as the Kellogg Youth Initiative Partnerships, Devolution, ENLACE (Engaging Latino Communities for Education), and the Native American Higher Education Initiative, to name just a few, has provided ample evidence of the effectiveness of these methods.

Learning and using tools like logic models can serve to increase the practitioner's voice in the domains of planning, design, implementation, analysis, and knowledge generation. The process of developing the model is an opportunity to chart the course. It is a conscious process that creates an explicit understanding of the challenges ahead, the resources available, and the timetable in which to hit the target. In addition, it helps keep a balanced focus on the big picture as well as the component parts.

In general, logic modeling can greatly enhance the participatory role and usefulness of evaluation as a management and learning tool. Developing and using logic models is an important step in building community capacity and strengthening community voice. The ability to identify outcomes and anticipate ways to measure them provides all program participants with a clear map of the road ahead. Map in hand, participants are more confident of their place in the scheme of things, and hence, more likely to actively engage and less likely to stray from the course – and when they do, to do so consciously and intentionally. Because it is particularly amenable to visual depictions, program logic modeling can be a strong tool in communicating with diverse audiences – those who have varying world views and different levels of experience with program development and evaluation.

Introduction

The *Logic Model Development Guide* contains four chapters and two comprehensive appendices.

Chapter 1 presents a basic introduction to the logic model as an action-oriented tool for program planning and evaluation. It also offers an array of sample logic models.

Chapter 2 consists of exercises and examples focused on the development of a simple program logic model. Exercises include practical examples, checklists for reviewing content quality, and a template for developing a logic model.

Chapter 3 gives instructions on how to expand a basic logic model to explore and explain the theory-of-change that describes the rationale for your program. A template and checklist are provided.

Chapter 4 offers two exercises that afford the reader with an introduction to how the basic logic modeling techniques introduced in the previous chapters can be applied to inform thinking about what should be included in an evaluation plan. Templates and checklists are also provided.

The ***Resources Appendix*** provides logic model development resources – references and Web sites worth visiting. The ***Forms Appendix*** includes blank templates to copy when developing your own logic models.

Acknowledgements

This work builds on the experience of many at the W.K. Kellogg Foundation who pioneered the application of logic modeling to their initiatives. For example, logic models were first used with the Kellogg Youth Initiative Partnerships (KYIP). In this application, the models were instrumental in helping staff establish program direction, implementation, an evaluation framework, and outcomes across three sites. In KYIP, logic modeling was used to facilitate and guide the development of the specific assumptions and processes that ultimately led to the transition of the initiative from a WKKF-operated program to a community-owned program. WKKF program staff, including Tyrone Baines, Phyllis Meadows, Gerald Smith, Judy Watson Olson, Steve Peffers, Joyce Brown, and John Seita were instrumental in these efforts.

Our work in developing the *Logic Model Development Guide* began at the request of Kellogg Foundation Program Director Blas Santos who expressed a need for user-friendly tools and processes to support the work of grantees in Latin America and the Caribbean.

The *Logic Model Development Guide* represents a collaborative effort. We particularly want to acknowledge the efforts of the Kellogg Foundation's former director of evaluation, Ricardo Millett, and his team of evaluation managers, including Astrid Hendricks-Smith and Mark Lelle, who have since left the organization. Their tireless work among staff and grantees continues to promote the use of logic models to plan, design, and manage initiatives. Dale Hopkins and Karin Ladley were instrumental in bringing the material to print. We also wish to acknowledge the work of the Kellogg Foundation Vice Presidents of Programs Rick Foster, Gail McClure, Dan Moore, and Gloria Smith, along with Senior Vice President of Programs Anne Petersen, who have underscored the importance of evaluation, embraced the logic model approach, and adopted it as a valued program support tool.

Special thanks are extended to Cynthia Phillips, a primary writer and consultant throughout the development of this guide, and Work Volk Consultants, LLP, for formatting and editorial assistance. Thanks, also, to Beverly Parsons of In Sites; Andrew Hahn and the students at the Florence Heller Graduate School for Advanced Studies in Social Welfare, Brandeis University; Marc Osten, Summit Consulting Collaborative; Sally Bond, The Program Evaluation Group; Joel Meister and Eva Moya, University of Arizona; Amy Coates-Madsen and staff at Maryland Association of Nonprofit Organizations; and Gail Randall, Greater Worchester Community Foundation.

–The Program Staff of the W.K. Kellogg Foundation

Chapter 1

Introduction to Logic Models

Chapter One defines logic models and explains their usefulness to program stakeholders. You will learn the relevance of this state-of-the-art tool to program planning, evaluation, and improvement.

Effective program evaluation does more than collect, analyze, and provide data. It makes it possible for you – program stakeholders – to gather and use information, to learn continually about and improve programs that you operate in or fund. The W.K. Kellogg Foundation believes evaluation – especially program logic model approaches – is a learning and management tool that can be used throughout a program's life – no matter what your stake in the program. Using evaluation and the logic model results in effective programming and offers greater learning opportunities, better documentation of outcomes, and shared knowledge about *what works* and *why*. The logic model is a beneficial evaluation tool that facilitates effective program planning, implementation, and evaluation.

A program logic model is a picture of how your program works – the theory and assumptions underlying the program. ...This model provides a road map of your program, highlighting how it is expected to work, what activities need to come before others, and how desired outcomes are achieved (p. 35).

W.K. Kellogg Foundation Evaluation Handbook (1998)

The *What* and *Why* of the Logic Model

The *WHAT*: Logic Model Definition

Basically, a logic model is a systematic and visual way to present and share your understanding of the relationships among the resources you have to operate your program, the activities you plan, and the changes or results you hope to achieve.

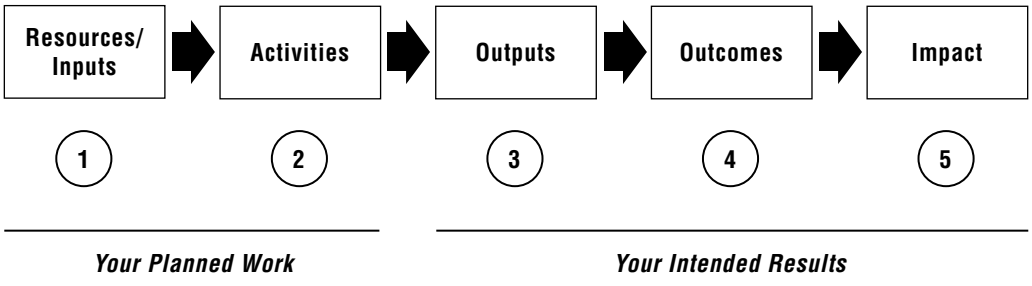


Figure 1. The Basic Logic Model.

The most basic logic model is a picture of how you believe your program will work. It uses words and/or pictures to describe the sequence of activities thought to bring about change and how these activities are linked to the results the program is expected to achieve.

Chapter 1

The Basic Logic Model components shown in Figure 1 above are defined below. These components illustrate the connection between *your planned work* and *your intended results*. They are depicted numerically by steps 1 through 5.

YOUR PLANNED WORK describes what resources you think you need to implement your program and what you intend to do.

Most of the value in a logic model is in the process of creating, validating, and modifying the model … The clarity of thinking that occurs from building the model is critical to the overall success of the program (p. 43).

W.K. Kellogg Foundation Handbook (1998)

1. Resources include the human, financial, organizational, and community resources a program has available to direct toward doing the work. Sometimes this component is referred to as ***Inputs***.

2. Program Activities are what the program does with the resources. **Activities** are the processes, tools, events, technology, and actions that are an intentional part of the program implementation. These interventions are used to bring about the intended program changes or results.

YOUR INTENDED RESULTS include all of the program's desired results (outputs, outcomes, and impact).

3. Outputs are the direct products of program activities and may include types, levels and targets of services to be delivered by the program.

4. Outcomes are the specific changes in program participants' behavior, knowledge, skills, status and level of functioning. Short-term outcomes should be attainable within 1 to 3 years, while longer-term outcomes should be achievable within a 4 to 6 year timeframe. The logical progression from short-term to long-term outcomes should be reflected in impact occurring within about 7 to 10 years.

5. Impact is the fundamental intended or unintended change occurring in organizations, communities or systems as a result of program activities within 7 to 10 years. In the current model of WKKF grantmaking and evaluation, impact often occurs after the conclusion of project funding.

The term *logic model* is frequently used interchangeably with the term *program theory* in the evaluation field. Logic models can alternatively be referred to as *theory* because they describe how a program works and to what end (definitions for each employed by leading evaluation experts are included in the Resources Appendix).

The *What*: How to "Read" a Logic Model

When "read" from left to right, logic models describe program basics over time from planning through results. Reading a logic model means following the chain of reasoning or "***If...then...***" statements which connect the program's parts. The figure below shows how the basic logic model is read.

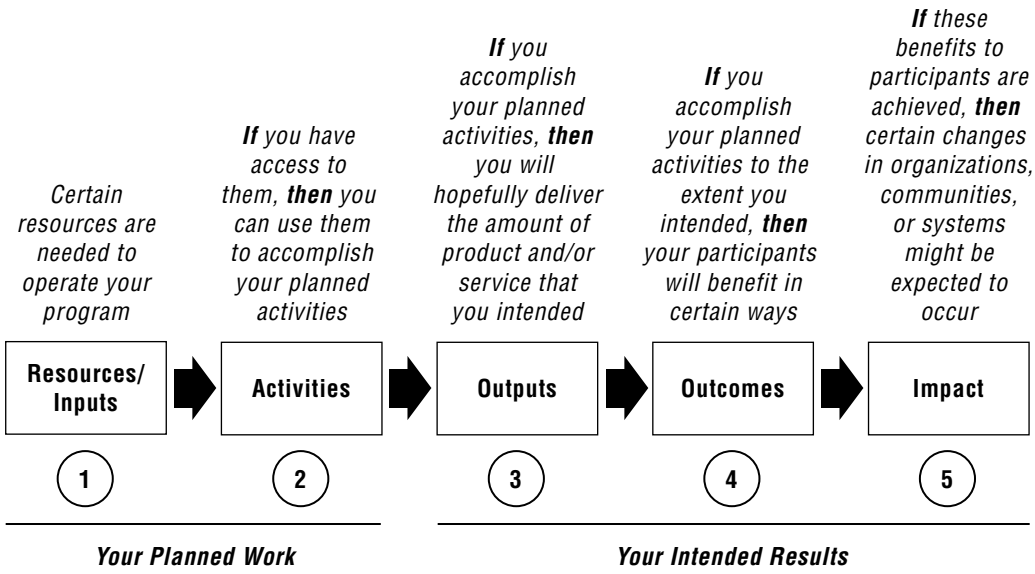


Figure 2. How to Read a Logic Model.

Sample Factors influencing the trip:

- *Family members' school and work schedules*
- *The holidays*
- *Winter weather*
- *Frequent Flier availability*

Sample Activities:

- *Creating/checking family schedules*
- *Gathering holiday flight and FF information*
- *Getting airport transportation*
- *Notifying Iowa relatives*

The *WHY*: Logic Model Purpose and Practical Application

The purpose of a logic model is to provide stakeholders with a road map describing the sequence of related events connecting the need for the planned program with the program's desired results. Mapping a proposed program helps you visualize and understand how human and financial investments can contribute to achieving your intended program goals and can lead to program improvements.

A logic model brings program concepts and dreams to life. It lets stakeholders try an idea on for size and apply theories to a model or picture of how the program would function. The following example shows how the logic model approach works. (If you are familiar with logic models, you may wish to skip ahead to the section entitled "Why Use A Logic Model?")

An Example:

We are proposing an inexpensive family trip from Charleston, South Carolina, to Des Moines, Iowa, to visit relatives during December school holidays. The seasonal trip we dream of taking from Charleston to Des Moines is the "program." Basic assumptions about our trip "program" are:

- We want to visit relatives between 12/10/00 and 1/5/01 while the children are out of school.
- We will fly from South Carolina to Iowa because it takes less time than driving and because frequent flier (FF) miles are available.
- Using frequent flier miles will reduce travel costs.

We have to determine the factors influencing our trip, including necessary resources, such as, the number of family members, scheduled vacation time, the number of frequent flier miles we have, round trip air reservations for each family member, and transportation to and from our home to the airport. The activities necessary to make this happen are the creation of our own family holiday schedule, securing our Iowa relative's schedule, garnering air line information and reservations and planning for transportation to and from the airport.

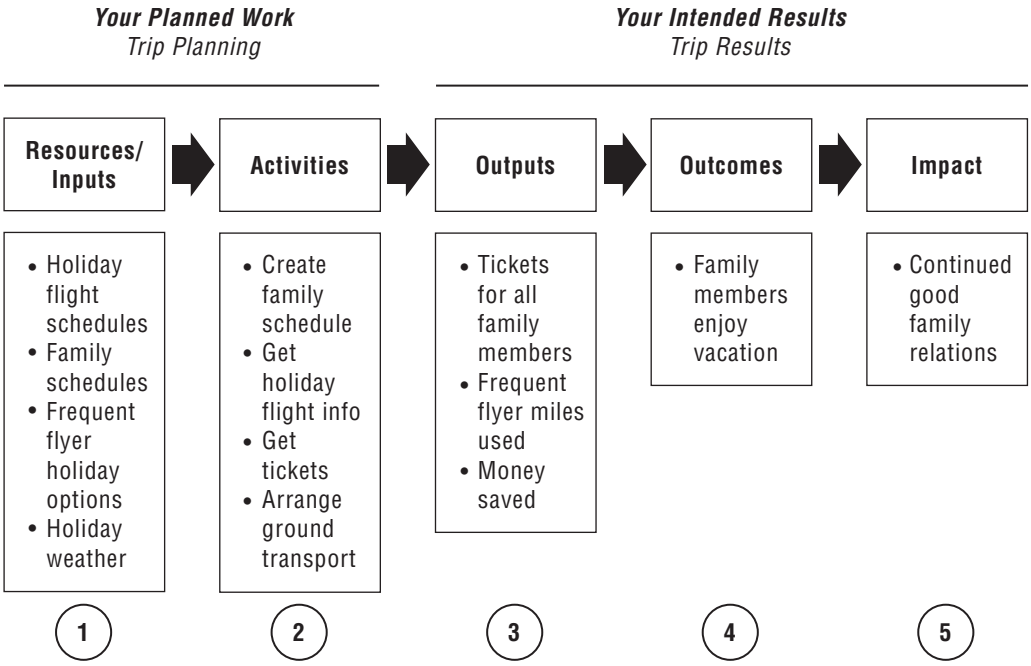
In this example, the results of our activities – or outputs – are mostly information, such as family schedules, flight schedules, and cost information based on the time frame of the trip. This information helps identify outcomes or immediate goals. For instance, if we make reservations as soon as possible, we are able to find flights with available frequent flier slots and probably have more options for flights that fit within the time frame. Knowing this, our outcomes improve – reservations made well in advance result in flight schedules and airline costs that suit our timeline and travel budget. Longer-term impact of our trip is not an issue here, but might be projected as continued good family relationships in 2010.

You can't do "good" evaluation if you have a poorly planned program.

Beverly Anderson Parsons (1999)

Using a simple logic model as a trip-planning tool produced tangible benefits. It helped us gather information to influence our decisions about resources and allowed us to meet our stated goals. Applying this process consistently throughout our trip planning positions us for success by laying out the best course of action and giving us benchmarks for measuring progress – when we touch down in Charlotte and change planes for Cincinnati, we know we're on course for Des Moines.

Typical logic models use table and flow chart formats like those presented here to catalogue program factors, activities, and results and to illustrate a program's dimensions. Most use text and arrows or a graphic representation of program ideas. This is what our trip planning "program" could look like in logic model format.



It was easy to organize travel plans in a flow chart, but we could also choose to organize and display our thinking in other ways. A logic model does not have to be linear. It may appear as a simple image or concept map to describe more complex program concepts. Settling on a single image of a program is sometimes the most difficult step for program stakeholders.

Why Use a Logic Model?

If program planners don't have any hypotheses guiding them, their potential for learning from the initiative is low, and the program is probably in trouble (p. 1).

Everything You Wanted to Know About Logic Models but Were Afraid to Ask,

Connie Schmitz and Beverly Anderson Parsons (1999)

As you can see from the travel plan example, logic models are useful tools in many ways. Because they are pictorial in nature, they require systematic thinking and planning to better describe programs. The visual representation of the master plan in a logic model is flexible, points out areas of strength and/or weakness, and allows stakeholders to run through many possible scenarios to find the best. In a logic model, you can adjust approaches and change courses as program plans are developed. Ongoing assessment, review, and corrections can produce better program design and a system to strategically monitor, manage, and report program outcomes throughout development and implementation.

Effective evaluation and program success rely on the fundamentals of clear stakeholder assumptions and expectations about how and why a program will solve a particular problem, generate new possibilities, and make the most of valuable assets. The logic model approach helps create shared understanding of and focus on program goals and methodology, relating activities to projected outcomes.

Logic Models Better Position Programs For Success

Many evaluation experts agree that use of the logic model is an effective way to ensure program success. Using a logic model throughout your program helps organize and systematize program planning, management, and evaluation functions.

The bane of evaluation is a poorly designed program.

Ricardo Millett, Director, WKKF Evaluation Unit

1. In ***Program Design and Planning,*** a logic model serves as a planning tool to develop program strategy and enhance your ability to clearly explain and illustrate program concepts and approach for key stakeholders, including funders.

Logic models can help craft structure and organization for program design and build in self-evaluation based on shared understanding of what is to take place. During the planning phase, developing a logic model requires stakeholders to examine best practice research and practitioner experience in light of the strategies and activities selected to achieve results.

2. In ***Program Implementation,*** a logic model forms the core for a focused management plan that helps you identify and collect the data needed to monitor and improve programming.

Using the logic model during program implementation and management requires you to focus energies on achieving and documenting results. Logic models help you to consider and prioritize the program aspects most critical for tracking and reporting and make adjustments as necessary.

3. For ***Program Evaluation and Strategic Reporting,*** a logic model presents program information and progress toward goals in ways that inform, advocate for a particular program approach, and teach program stakeholders.

Chapter 1

We all know the importance of reporting results to funders and to community stakeholders alike. Communication is a key component of a program's success and sustainability. Logic models can help strategic marketing efforts in three primary ways:

- ***Describing programs*** in language clear and specific enough to be understood and evaluated.
- ***Focusing attention and resources*** on priority program operations and key results for the purposes of learning and program improvement.
- ***Developing targeted communication*** and marketing strategies.

There are many ways to conduct evaluations, and professional evaluators tend to agree that there is no "one best way" to do any evaluation. Instead, good evaluation requires carefully thinking through the questions that need to be answered, the type of program being evaluated, and the ways in which the information generated will be used. Good evaluation, in our view, should provide useful information about program functioning that can contribute to program improvement.

W.K. Kellogg Foundation Evaluation Unit

The Table below describes the relationship between a successful program and the benefits derived from the use of logic models.

Program Elements	Criteria for Program Success[1]	Benefits of Program Logic Models[2]
Planning and Design	Program goals and objectives, and important side effects are well defined ahead of time.	Finds "gaps" in the theory or logic of a program and work to resolve them.
	Program goals and objectives are both plausible and possible.	Builds a shared understanding of what the program is all about and how the parts work together.
Program Implementation and Management	Relevant, credible, and useful performance data can be obtained.	Focuses attention of management on the most important connections between action and results.
Evaluation, Communication, and Marketing	The intended users of the evaluation results have agreed on how they will use the information.	Provides a way to involve and engage stakeholders in the design, processes, and use of evaluation.

How Logic Models Better Position Programs Toward Success.

Logic Models Strengthen the Case for Program Investment

Clear ideas about what you plan to do and why – as well as an organized approach to capturing, documenting, and disseminating program results – enhance the case for investment in your program.

[1] Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (Eds.). (1994). *Handbook of Practical Program Evaluation.* San Francisco: Jossey-Bass Publishers.

[2] Barley, Z., Phillips, C., & Jenness, M. (1998). *Decoding Program Logic Models.* Workshop presented at the Annual Meeting of the American Evaluation Association, Chicago, IL, November, 1998.

Developing a Program Logic Model Requires a Simple Image and a Straightforward Approach

A picture IS worth a thousand words. The point of developing a logic model is to come up with a relatively simple image that reflects how and why your program will work. Doing this as a group brings the power of consensus and group examination of values and beliefs about change processes and program results.

LOGIC MODEL
IF… THEN
Assumptions:

- Certain resources are needed to operate your program.
- ***If*** you have access to them, ***then*** you can use them to accomplish your planned activities.
- ***If*** you accomplish your planned activities, ***then,*** you will, it is hoped, deliver the amount of product and/or service that you intended.
- ***If*** you accomplish your planned activities to the extent intended, ***then*** your participants will benefit in specific ways.
- ***If*** these benefits to participants are achieved, ***then*** certain changes in organizations, communities, or systems might occur under specified conditions.

Logic Models Reflect Group Process and Shared Understanding

Frequently, a professional evaluator is charged with developing a logic model for program practitioners. But a logic model developed by all stakeholders – program staff, participants, and evaluators – produces a more useful tool and refines program concepts and plans in the process. We recommend that a logic model be developed collaboratively in an inclusive, collegial process that engages as many key stakeholders as possible. This guide provides a step-by-step process to assist program planners.

Like Programs, Logic Models Can Change Over Time

As a program grows and develops, so does its logic model. A program logic model is merely a snapshot of a program at one point in time; it is not the program with its actual flow of events and outcomes. A logic model is a work in progress, a working draft that can be refined as the program develops.

Simple Logic Model Basics

Creating a logic model: What they look like and what needs to be included

Logic models come in as many sizes and shapes as the programs they represent. A simple model focuses on project-level results and explains five basic program components. The elements outlined below are typical of the model promoted by United Way of America to support an outcomes-based approach to program planning and evaluation.

Developing and Reading a Basic Logic Model

Read from left to right, logic models describe program basics over time, beginning with best practice information or knowledge about "what works" from successful program practitioners and other trusted authorities. Reading a logic model means following the chain of reasoning or "***If...then...***" statements which connect the program's parts. The gray box in the left column defines the assumptions stated in "***If...then...***" terms.

Chapter 1

Building a Logic Model by Basic Program Components

As you conceptualize your program, begin by describing your basic assumptions and then add the following program components in the order that they should occur.

1. ***Factors*** are resources and/or barriers, which potentially enable or limit program effectiveness. Enabling *protective factors* or *resources* may include funding, existing organizations, potential collaborating partners, existing organizational or interpersonal networks, staff and volunteers, time, facilities, equipment, and supplies. Limiting *risk factors* or *barriers* might include such things as attitudes, lack of resources, policies, laws, regulations, and geography.

2. ***Activities*** are the processes, techniques, tools, events, technology, and actions of the planned program. These may include *products* – promotional materials and educational curricula; *services* – education and training, counseling, or health screening; and *infrastructure* – structure, relationships, and capacity used to bring about the desired results.

3. ***Outputs*** are the *direct results* of program activities. They are usually described in terms of the *size and/or scope of the services and products delivered or produced* by the program. They indicate if a program was delivered to the intended audiences at the intended "dose." A program output, for example, might be the *number* of classes taught, meetings held, or materials produced and distributed; program *participation rates* and demography; or *hours of each type of service* provided.

4. ***Outcomes*** are specific *changes in attitudes, behaviors, knowledge, skills, status, or level of functioning* expected to result from program activities and which are most often expressed *at an individual level*.

5. ***Impacts*** are *organizational, community, and/or system level changes* expected to result from program activities, which might include improved conditions, increased capacity, and/or changes in the policy arena.

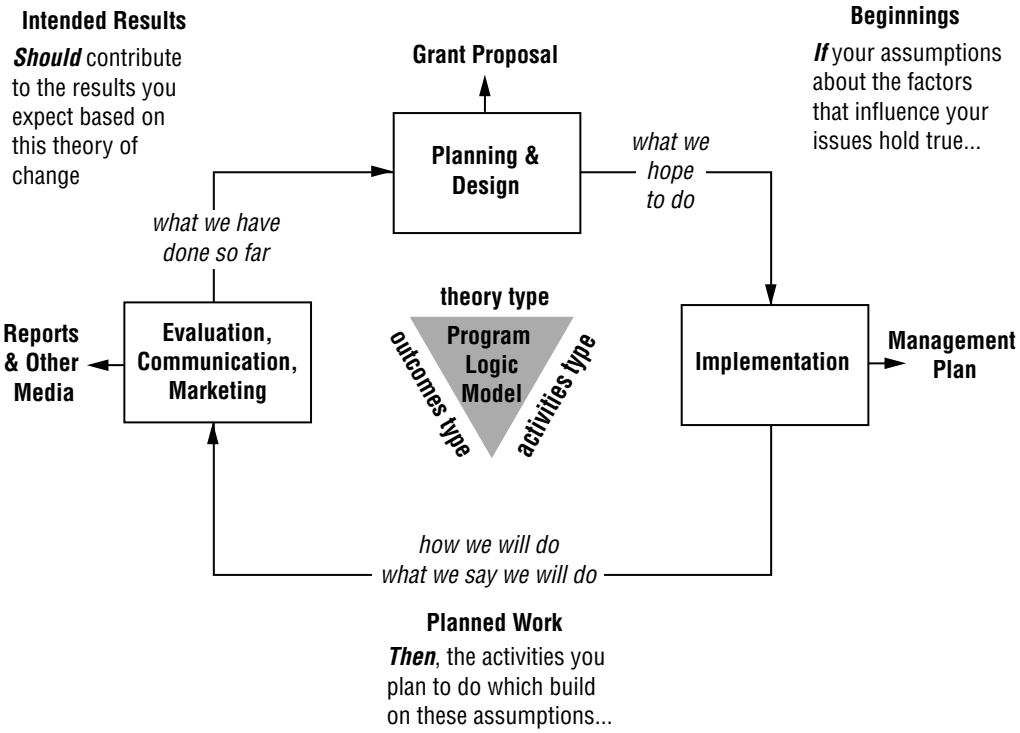
Thinking about a program in logic model terms prompts the clarity and specificity required for success, and often demanded by funders and your community. Using a simple logic model produces (1) an inventory of what you have and what you need to operate your program; (2) a strong case for how and why your program will produce your desired results; and (3) a method for program management and assessment.

Other Logic Model Examples

In practice, most logic models are more complex and fall into one of three categories: the theory approach model (conceptual), outcome approach model, or activities approach model (applied) – or a blend of several types. It is not unusual for a program to use all three types of logic models for different purposes. No one model fits all needs, so you will

need to decide exactly what you want to achieve with your logic model – and where you are in the life of your program – before deciding on which model to use.

Types of Logic Models: Emphasis and Strengths



Types of Logic Models: Emphasis and Strengths
A program is a theory and an evaluation is its test. In order to organize the evaluation to provide a responsible test, the evaluator needs to understand the theoretical premises on which the program is based (p. 55).

Carol Weiss (1998)

Descriptions of Three Approaches to Logic Models: Which Fits Your Program?

1. ***Theory Approach Models*** emphasize the theory of change that has influenced the design and plan for the program. These logic models provide rich explanation of the reasons for beginning to explore an idea for a given program. Sometimes they have additional parts that specify the problem or issue addressed by the program, describe the reasons for selecting certain types of solution strategies, connect proven strategies to potential activities, and other assumptions the planners hold that influence effectiveness. These models illustrate how and why you think your program will work. They are built from the "big picture" kinds of thoughts and ideas that went into conceptualizing your program. They are coming to be most often used to make the case in grant proposals. Models describing the beginnings of a program in detail are most useful during program planning and design.

2. ***Outcomes Approach Models*** focus on the early aspects of program planning and attempt to connect the resources and/or activities with the desired results in a workable program. These models often subdivide outcomes and impact over time to describe short-term (1 to 3 years), long-term (4 to 6 years), and impact (7 to 10 years) that may result from a given set of activities. Although these models are developed with a theory of change in mind, this aspect is not usually emphasized explicitly. Models that outline the approach and expectations behind a program's intended results are most useful in designing effective evaluation and reporting strategies.

3. ***Activities Approach Models*** pay the most attention to the specifics of the implementation process. A logic model of this type links the various planned activities together in a manner that maps the process of program implementation. These models describe what a program intends to do and as such are most useful for the purposes of program monitoring and management. This type provides the detailed steps you think you will need to follow to implement your program. It shows what you will actually *do* in your community if your proposal is funded. Models that emphasize a program's planned work are most often used to inform management planning activities.

The purpose of using program logic models in WKKF grantmaking is to help internal and external stakeholders understand how the Foundation's investment will contribute to achieving the intended goals. This understanding should help these various stakeholders make informed decisions about program priorities, funding priorities, assistance to grantees, evaluation of programming impact, and marketing, communication, and marketing strategies.

W.K. Kellogg Foundation Evaluation Handbook (1998)

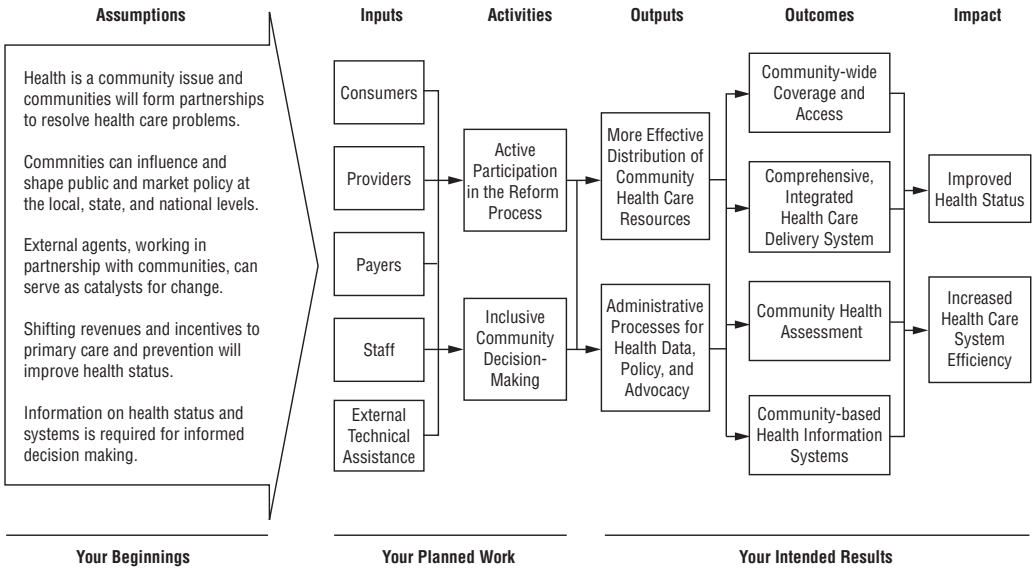
Working Through Theory Approach Logic Models Emphasizes Assumptions

A theory approach logic model links theoretical ideas together to explain underlying program ***assumptions***. The focus here is on the problem or issue and the reasons for proposing the solution suggested in your program's approach. Remember, the theory logic model is broad and about "big ideas," not about specific program "nuts and bolts."

Noted evaluator and program theorist Carol Weiss (1998) explains that for program planning, monitoring, and evaluation, it is important to know not only ***what*** the program expects to achieve but also ***how.*** We must understand the principles on which a program is based, a notion not included in evaluation until recently. Discussions about the *whethers*, *hows*, and *whys* of program success require credible evidence and attention to the paths by which outcomes and impacts are produced.

The theory logic model is suitable for use by funders and grantees. A case example of its use is provided below.

In this case, the model describes a WKKF cluster initiative's (Comprehensive Community Health Models of Michigan) programming strategy or its theory of change. Notice that this model places emphasis on "Your Beginnings" by including the assumptions identified by program planners as the principles behind the design of the initiative.



Example of a Theory Logic model (Adapted from WKKF's Comprehensive Community Health Models of Michigan).

These models help build a common understanding between managers and evaluators.... Such agreement is a prerequisite for evaluation work that is likely to be useful to management. [These models] display the key events (inputs, activities, outcomes) that could be monitored and the assumed causal linkages that could be tested in evaluations of the program.

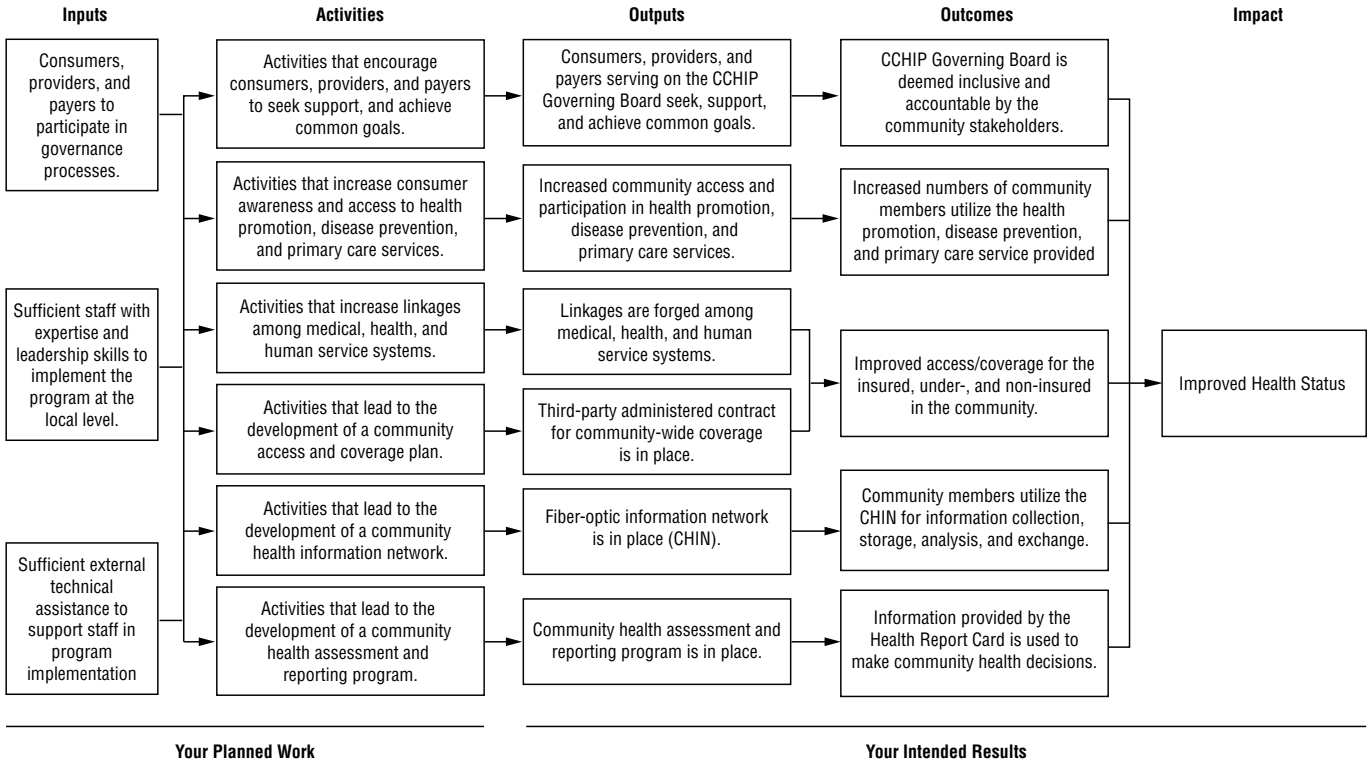
Joseph S. Wholey, Harry P. Hatry, and K.E. Newcomer (1994)

Working with Outcome Approach Models Highlights Activities and Program Implementation

Outcome approach logic models display the interrelationships between specific program activities and their outcomes. On the next page is an example drawn from the Calhoun County Health Improvement Program, funded under the Comprehensive Community Health Models of Michigan initiative.

This linear, columnar model emphasizes the ***causal linkages*** thought to exist among program components. The arrows show which sets of activities program developers believed would contribute to what outcomes. These statements serve as logical assertions about the perceived relationship among program operations and desired results and are the hallmark of the logic model process.

Notice that this model emphasizes "Your Intended Results" in the greatest relative detail and anticipates achievement outside the time allotted for the initiative.



Example of an Outcome Approach model (example drawn from the Calhoun County Health Improvement Program, funded under the Comprehensive Community Health Models of Michigan initiative).

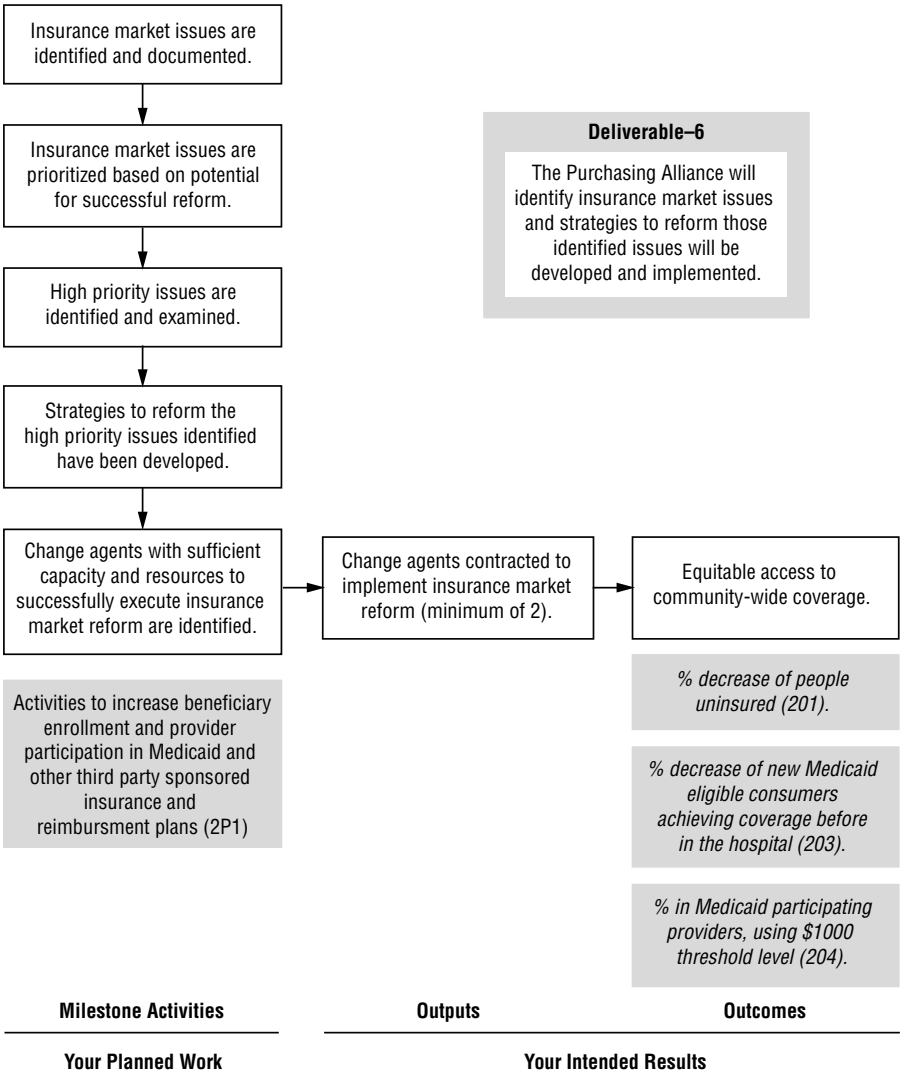
Not only will a logic model clarify each element of your program, it will enable you to respond to the question: "To what do I want to be held accountable?".

The Evaluation Forum (1999)

Using the Activities Approach Models to Track Outcomes

The activities approach logic model also connects program resources and activities to desired results but does so in very great detail. Each outcome is usually dealt with separately by the activities and events that must take place to keep the program on track. The model emphasizing "Your Planned Work" can be used as a work plan or management tool for program components and in conjunction with other models.

Notice how it points out what program activities need to be monitored and what kind of measurements might indicate progress toward results. Below is one model describing the connections between project tasks and outcome achievement for the community coverage strand from the outcome approach example provided earlier.



Adapted from the Calhoun County Health Improvement Program, one site of WKKF's Comprehensive Community Health Models of Michigan initiative

There Is No *Best* Logic Model

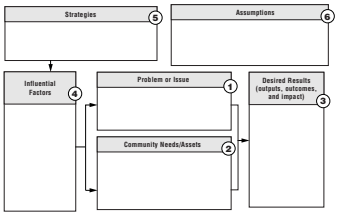
Try several on for size. Choose the model that fits your program best and provides the information you need in the format that is most helpful. Like anything else, it takes practice to use logic models as effective program tools. We learn through trial and error to find what works best for what program. Don't hesitate to experiment with program logic model design to determine what works best for your program. And don't be concerned if your model doesn't look like one of the case examples.

The following show how the logic model forms gather information that can be used throughout your program's life – from defining the theory on which your program rests to evaluating program impact.

Chapter 1

How to use a Logic Model Through the Life of Your Program:

1. Program Planning



For more detail, see the Program Planning Template on p. 57.

2. Program Implementation

RESOURCES	ACTIVITIES	OUTPUTS SHORT	SHORT & LONG-TERM OUTCOMES	IMPACT
In order to accomplish our set of activities we will need the following:	*In order to address our problem or asset we will accomplish the following activities:*	*We expect that once accomplished these activities will produce the following evidence or service delivery:*	*We expect that if accomplished these activities will lead to the following changes in 1-3 then 4-6 years:*	*We expect that if accomplished these activities will lead to the following changes in 7-10 years:*

For more detail, see the Program Implementation Template on p. 54.

3. Program Evaluation

Evaluation Focus Area	Audience	Question	Use

For more detail, see the Evaluation Planning Template on p. 59.

Focus Area	Question	Indicators	Technical Assistance Needed

For more detail, see the Indicators Development Template on p. 61.

CLARIFYING PROGRAM THEORY:

1. **PROBLEM OR ISSUE STATEMENT:** Describe the problem(s) your program is attempting to solve or the issue(s) your program will address.
2. **COMMUNITY NEEDS/ASSETS:** Specify the needs and/or assets of your community that led your organization to design a program that addresses the problem.
3. **DESIRED RESULTS (OUTPUTS, OUTCOMES AND IMPACTS):** Identify desired results, or vision of the future, by describing what you expect to achieve near- and long-term.
4. **INFLUENTIAL FACTORS:** List the factors you believe will influence change in your community.
5. **STRATEGIES:** List general successful strategies or "best practices" that have helped communities like yours achieve the kinds of results your program promises.
6. **ASSUMPTIONS:** State the assumptions behind *how* and *why* the change strategies will work in your community.

DEMONSTRATING YOUR PROGRAM'S PROGRESS:

1. **OUTPUTS:** For each program activity, identify what outputs (service delivery/implementation targets) you aim to produce.
2. **OUTCOMES:** Identify the short-term and long-term outcomes you expect to achieve for each activity.
3. **IMPACT:** Describe the impact you anticipate in your community in 7 to 10 years with each activity as a result of your program.
4. **ACTIVITIES:** Describe each of the activities you plan to conduct in your program.
5. **RESOURCES:** Describe the resources or influential factors available to support your program activities.

PROGRAM EVALUATION QUESTIONS AND INDICATORS:

1. **FOCUS AREA:** From your program theory logic model, list the components of the most important aspects of your program.
2. **AUDIENCE:** Identify the key audiences for each focus area. Who has an interest in your program?
3. **QUESTIONS:** For each focus area and audience, list the questions they may have about your program.
4. **INFORMATION USE:** For each audience and question you have identified, identify the ways you will use the evaluation information.
5. **INDICATORS:** Describe what information could be collected that would indicate the status of your program and its participants for each question.
6. **TECHNICAL ASSISTANCE:** Indicate the extent to which your organization has the evaluation and data management expertise to collect and analyze the data that relates to this indicator.

Chapter 2

Developing a Basic Logic Model For Your Program

Drawing a picture of how your program will achieve results

Whether you are a grantseeker developing a proposal for start-up funds or a grantee with a program already in operation, developing a logic model can strengthen your program. Logic models help identify the factors that will affect your program and enable you to anticipate the data and resources you will need to achieve success. As you engage in the process of creating your program logic model, your organization will systematically address these important program planning and evaluation issues:

Over the past few years, I have markedly changed my approach to logic modeling. I have become convinced that it makes a considerable difference if you do the outcomes before planning the activities.

I definitely advocate doing the outcomes first! I find that people come up with much more effective activities when they do. Use the motto, "plan backward, implement forward."

Beverly Anderson Parsons, WKKF Cluster Evaluator

- Cataloguing of the resources and actions you believe you will need to reach intended results.
- Documentation of connections among your available resources, planned activities and the results you expect to achieve.
- Description of the results you are aiming for in terms of specific, measurable, action-oriented, realistic and timed outcomes.

The exercises in this chapter gather the raw material you need to draw a basic logic model that illustrates how and why your program will work *and* what it will accomplish. You can benefit from creating a logic model at any point in the life of any program. The logic model development process helps people inside and outside your organization understand and improve the purpose and process of your work.

Chapter 2 is organized into two sections – Program Implementation, and Program Results. The best recipe for program success is to complete both exercises. (Full-size masters of each exercise and the checklists are provided in the Forms Appendix at the back of the guide for you to photocopy and use with stakeholder groups as you design your program.)

Exercise 1: Program Results. In a series of three steps, you describe the results you plan to achieve with your program.

Exercise 2: Program Resources and Activities by taking you through three steps that connect the program's resources to the actual activities you plan to do.

The Mytown Example
Throughout Exercises 1 and 2 we'll follow an example program to see how the logic model steps can be applied. In our example, the folks in Mytown, USA, are striving to meet the needs of growing numbers of uninsured residents who are turning to Memorial Hospital's Emergency Room for care. Because that care is expensive and not the best way to offer care, the community is working to create a free clinic. Throughout the chapters, Mytown's program information will be dropped into logic model templates for Program Planning, Implementation, and Evaluation.

Chapter 2

Novice logic modelers may want to have copies of the Basic Logic Model Template in front of them and follow along. Those readers with more experience and familiarity may want to explore the text and then skip ahead to the completed Basic Logic Model for the Mytown Example on page 34.

Demonstrating Progress Toward Change

The Importance of Documenting Progress

According to many funders, grant applications frequently lack solid descriptions of how programs will demonstrate their effectiveness. Some grantees think activities are ends unto themselves. They report the numbers of participants they reach or the numbers of training sessions held as though they were results.

Conducting an activity is *not* the same as achieving results from the accomplishment of that activity. For example, being seen by a doctor is different from reducing the number of uninsured emergency room visits. Tracking data like meetings held or patients enrolled *does* monitor your program's implementation and performance, but those data are outputs (activity data), not outcomes (which refer to the results you expect to achieve in future years).

"Do the outcomes first" is sage advice. Most logic models lack specific short- and long-term outcomes that predict what will be achieved several years down the road. Specifying program milestones *as you design the program* builds in ways to gather the data required and allows you to periodically assess the program's progress toward the goals you identify. **For that reason, Exercise 1 isn't filled out from left to right. This exercise asks you to "do the outcomes first." We will focus our attention first on what we have called "*your intended results.*"**

As you implement your program, outcome measures enhance program success by assessing your progress from the beginning and all along the way. That makes it possible to notice problems early on. The elements (Outputs, Outcomes, and Impact) that comprise *your intended results* give you an outline of what is most important to monitor and gauge to determine the effectiveness of your program. You can correct and revise based on your interpretation of the collected data.

Exercise 1 – Describing Results

Describe the results you desire – Outputs, Outcomes and Impact

If you were running the Mytown Free Clinic, how would you show that your desired outcome (a reduction in uninsured emergency care) didn't result from a mass exodus of uninsured residents from Mytown, USA, or a sudden increase in number of employees offered health insurance coverage by local businesses?

How will you demonstrate that *your program* contributed to the change you intend? A well-crafted logic model can assert it is reasonable to claim that your program made a substantive contribution to your intended change. When programs operate in real communities where influences and forces are beyond your control, evaluation is generally more about documenting a program's contribution than about proving something. Community-based initiatives operate in complex environments where the scientific certainty of "proof" is seldom attainable. This is where logic models can be especially helpful.

INSTRUCTIONS: Exercise 1 will use the Basic Logic Model Development Template. In particular, you will use the information presented in the gray text boxes that follow about the Mytown example program to determine what results are intended for this program. Example information about outcomes, impacts, and outputs are provided. You will fill in the blank Basic Logic Model Development Template to illustrate first the outcomes and impacts sought and then the outputs. You can then look at the completed template on page 25 to see compare your interpretation with that produced by the Mytown folks.

Exercise 1 uses the Basic Logic Model Development Template

Resources	Activities	Outputs	Short- & Long-Term Outcomes	Impact
In order to accomplish our set of activities we will need the following:	*In order to address our problem or asset we will conduct the following activities:*	*We expect that once completed or under way these activities will produce the following evidence of service delivery:*	*We expect that if completed or ongoing these activities will lead to the following changes in 1–3 then 4–6 years:*	*We expect that if completed these activities will lead to the following changes in 7–10 years:*

Outcomes and Impacts should be SMART:

- Specific
- Measurable
- Action-oriented
- Realistic
- Timed

Some logic models number the lists within a column to aid discussion. Some tabular logic models use rows to order and show the relationships among components. Some logic models, like the outcome and activity examples provided in Chapter One, use a box and arrow format to illustrate the "causal linkages" demonstrating how your resources, activities, outputs, outcomes, and impact connect to form chains. These depictions add to the clarity of your logic model/evaluation plan. However, for the most basic of logic models, the inventory approach we illustrate is sufficient to capture your thinking about how a program will work. The other techniques will improve its utility, but the most important task is to first get the component parts categorized and described. Once you have completed the inventory table for this and Exercise 2 feel free to experiment with identifying the relationships among the items across columns.

Short-term outcomes are results you expect to achieve one to three years after a program activity is under way.

Short-term outcomes are specific changes in things like attitudes, behaviors, knowledge, skills, status, or level of functioning expected to result from program activities. These usually are expressed at an individual level among program participants.

EXAMPLES: Signed Memorandum of Agreement from the local technical college donating clinic space, change in participants' attitudes about the need for a medical home, increase in numbers of scheduled annual physicals, increased patient follow-up visits, change in staff's awareness of patient scheduling challenges, increased appropriate referrals from ER's.

Insert Mytown's short-term outcomes in the Short- and Long-Term Outcomes Column of the Basic Logic Model Development Template.

Long-term outcomes are results you expect to achieve in four to six years.

Long-term outcomes are also specific changes in things like attitudes, behaviors, knowledge, skills, status, or level of functioning expected to result from program activities. These usually build on the progress expected by the short-term outcomes.

EXAMPLES: The clinic serves as a medical home for 500 uninsured patients. The clinic has sustained funding sources: patient co-payments ($10/visit) provide 20% of the Clinic's operating costs, United Way provides 20%, Memorial Hospital donates 20%, the Medical Society contributes 20% and an endowment established at the Community Foundation provides the final 20%. An annual golf tournament organized by the Kiwanis Club funds special clinic projects. There has been a 25% reduction in uninsured emergency care since Mytown Free Clinic opened five years ago. In the Clinic's fifth year there is a 15% reduction in uninsured ER visits. Seventy-five volunteer administrators and 300 volunteer medical professionals regularly serve at the clinic each year. Five companies donate all necessary medical supplies. Grant funds purchase the computers and software needed to create electronic patient records. For five years patient satisfaction ratings have been 90%.

Insert Mytown's long-term outcomes in the Short- and Long-Term Outcomes column of the Basic Logic Model Development Template.

Impact refers to the results expected seven to ten years after an activity is under way – the future social change your program is working to create.

Impacts are the kinds of organizational, community, or system level changes expected to result from program activities and which might include improved conditions, increased capacity, and/or changes in the policy arena.

EXAMPLES: Specific reduction in inappropriate emergency room use, increased donations of clinic supplies to meet identified needs, a stable supply of medical volunteers, an endowment supporting 35% of the clinic's operating funds, 900 patients served/year.

Insert Mytown's impacts in the Impact Column of the Basic Logic Model Development Template.

Outputs are data about activities.

They are the direct results of program activities. They are usually described in terms of size and scope of the services or products delivered or produced by the program. They indicate whether or not a program was delivered to the intended audiences at the intended "dose." A program output, for example, might include the number of classes taught, meetings held, materials distributed, program participation rates, or total service delivery hours.

EXAMPLES: Number of patients referred to the Free Clinic from Memorial ER/year, the number of patients screened/year, the number of qualified patients enrolled in the Free Clinic/year, the average number of patient visits/day, the total number of patient visits/year, the number and specialties of medical volunteers, the number of volunteer administrators trained, the number and locations of clinic posters distributed, the number of potential patients calling for information/ month.

Insert Mytown's outputs in the Outputs Column of the Basic Logic Model Development Template.

Exercise 1 Checklist:

Review what you have created using the checklist below to assess the quality of your draft.

	Progress Toward Results Quality Criteria	Yes	Not Yet	Comments Revisions
1.	A variety of audiences are taken into consideration when specifying credible outputs, outcomes, and impacts.	☐	☐	
2.	Target participants and/or partners are described and quantified as outputs (e.g. 100 teachers from 5 rural high schools).	☐	☐	
3.	Events, products, or services listed are described as outputs in terms of a treatment or dose (e.g. 30 farmers will participate in at least 3 sessions of program, or curriculum will be distributed to at least 12 agencies).	☐	☐	
4.	The intensity of the intervention or treatment is appropriate for the type of participant targeted (e.g. higher-risk participants warrant higher intensities).	☐	☐	
5.	The duration of the intervention or treatment is appropriate for the type of participant targeted (e.g. higher-risk participants warrant longer duration).	☐	☐	
6.	Outcomes reflect reasonable, progressive steps that participants can make toward longer-term results.	☐	☐	
7.	Outcomes address awareness, attitudes, perceptions, knowledge, skills, and/ or behavior of participants.	☐	☐	
8.	Outcomes are within the scope of the program's control or sphere of reasonable influence.	☐	☐	
9.	It seems fair or reasonable to hold the program accountable for the outcomes specified.	☐	☐	
10.	The outcomes are specific, measurable, action-oriented, realistic, and timed.	☐	☐	
11.	The outcomes are written as change statements (e.g. things increase, decrease, or stay the same).	☐	☐	
12.	The outcomes are achievable within the funding and reporting periods specified.	☐	☐	
13.	The impact, as specified, is not beyond the scope of the program to achieve.	☐	☐	

Exercise 2 – Describing Actions

Linking It All Together

Exercise 2 illustrates exactly how you plan to put your program theory to work. It leads you to identify the resources and activities your program will need to achieve your intended results. This exercise documents your knowledge of the community resources you have available and specific activities your program will implement.

I would emphasize that people may well change their minds about the activities that are the most useful after having done the results work.

Beverly Anderson Parsons, WKKF Cluster Evaluator

Program rationales in grant proposals are usually strong. Grantees tend to have a very good sense of *what* they want to do. However, they frequently fail to make specific connections between their program and related best practice literature and practitioner wisdom that could *and should* support their approach and their work.

To connect actions to program results, this exercise links your knowledge of what works with specific descriptions of what your program will do. It requires you to anticipate what will be needed to support program activities. The elements that comprise your program implementation act as a game plan for the program you propose.

Most logic models list activity items and resources (like planning meetings, curriculum purchase or design, training workshops, and service delivery). Depending on the nature of your effort, other types of products and processes may be included. Management-oriented logic models also include program and evaluation development, staff and volunteer training, recruitment of partners and participants, and the publicity needed to support your work along the way.

As mentioned earlier, if your program addresses multiple issues you may find it helpful to go through the exercises for each issue in turn and then aggregate them into a larger model that highlights the relationships among issues.

We recommend referring to a literature review on the problem your program is designed to address when you specify program activities. From this explicit knowledge of what works, you can more clearly connect the abstract strategies supporting the program to its concrete activities.

When Exercise 2 is complete and you are satisfied that you have an accurate inventory of the Mytown program's component parts, transfer the information to the **Basic Logic Model Development Template.** Remember you have already filled in the three columns on the right with what you have learned about the intended results for the Mytown program example.

What activities are planned? Based on what you know about effective ways to solve problems or build assets, what specific activities have you planned?

EXAMPLES: Personnel Committee launches and completes search for full-time director. Director is hired and oriented to the board and the community. Board and staff visit the Anywhere Free Clinic to learn from its experience and to select documents to replicate (i.e., policies and procedures, job descriptions, equipment needs, budgets, funding strategies, volunteer and patient records). Board and staff conduct program-planning retreat. Based upon Anywhere's funding plan, board secures Free Clinic's first-year funding. Marketing Committee creates public relations campaign in collaboration with Volunteer Committee to secure volunteers and patients. Facility Committee creates and completes MOA with technical college to secure a clinic facility. Quality Assurance Committee creates evaluation plan in cooperation with Memorial Hospital's Emergency Room staff and the local Chamber of Commerce.

Summarize Mytown's activities in the Activities column of the Basic Logic Model Development Template

What resources are needed? Once you have specified what you plan to do, determine the resources you will need to support the solutions your program proposes. For some types of programs, it may also be helpful to describe the influential factors you are counting on to support your efforts in the community.

EXAMPLES: Medical Society/Memorial Hospital Task Force for the Uninsured will become a Free Clinic Board of Directors and secure a 501(c)(3) status from the IRS. The Board will recruit 7–10 additional representatives from drug companies, the local technical school, Mytown's United Way, the Chamber of Commerce, the Community Foundation, the Volunteer Center, the Nurses Association, etc. During a 6-month planning period, board committees will be launched; staff will be recruited/hired/oriented; a site visit will be conducted; and the Clinic's first-year's funding ($150,000/year) will be secured. Committees will create an MOA with Memorial Hospital and the Medical Society to secure equipment required: 5 exam tables, 7 desks, 5 blood pressure cuffs, 5 otoscopes, 5 stethoscopes, 5 PDR's, 1 set of scales, 10 thermometers, three computers, one first aid emergency kit.

Summarize Mytown's resources in the Resources column of the Basic Logic Model Development Template.

Exercise 2 Checklist:

Review what you have created using the checklist below to assess the quality of your draft.

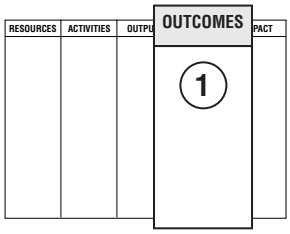
Theory into Action Quality Criteria	Yes	Not Yet	Comments/Revisions
1. Major activities needed to implement the program are listed.	☐	☐	
2. Activities are clearly connected to the specified program theory.	☐	☐	
3. Major resources needed to implement the program are listed.	☐	☐	
4. Resources match the type of program.			
5. All activities have sufficient and appropriate resources.	☐	☐	

Chapter 2

Here we include a flowchart that summarizes the steps to complete your basic logic model. Keep in mind that you could use this inventory style template to then further describe the relationships among the components using numbered items, rows, or boxes and arrows as we mentioned earlier.

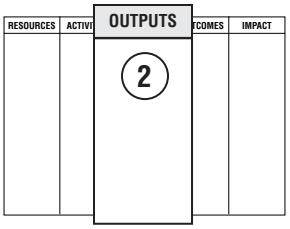
Flowchart for Exercises 1 & 2 – Describing Results, Resources, and Activities

Exercise 1 Describing Results



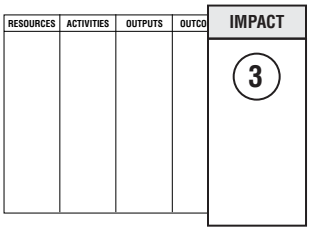
Step 1.1

For each of the specific activities you have planned to do, what short-term and then long-term outcomes do you expect to achieve as indicators of the progress made by your program toward its desired results?



Step 1.2

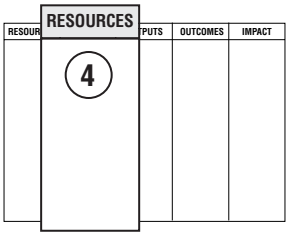
For each of the specific activities that you have planned to do, what outputs (service delivery or implementation targets) do you hope to reach through the operation of your program?



Step 1.3

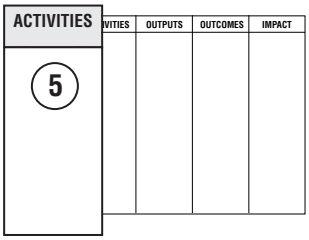
For each of the specific activities you have planned to do, what impact do you expect to achieve in your community?

Exercise 2 Describing Resources and Activities



Step 2.1

Knowing what you know about what works to solve problems or build assets as specified in the theory of change for your program, what specific activities have you planned to do?



Step 2.2

What resources are available to your program to support the specific activities you have planned to do (for some programs, it may also be important to state those influential factors you are counting on to support your work)?

Logic Model Development
Program Implementation Template – Exercise 1 & 2

RESOURCES	ACTIVITIES	OUTPUTS	SHORT- AND LONG-TERM OUTCOMES	IMPACT
In order to accomplish our set of activities we will need the following:	*In order to address our problem or asset we will accomplish the following activities:*	*We expect that once accomplished these activities will produce the following evidence or service delivery:*	*We expect that if accomplished these activities will lead to the following changes in 1–3 then 4–6 years:*	*We expect that if accomplished these activities will lead to the following changes in 7–10 years:*
• IRS 501(c)(3) status • Diverse, dedicated board of directors representing potential partners • Endorsement from Memorial Hospital, Mytown Medical Society, and United Way • Donated clinic facility • Job descriptions for board and staff • First year's funding ($150,000) • Clinic equipment • Board & staff orientation process • Clinic budget	• Launch/complete search for executive director • Board & staff conduct Anywhere Free Clinic site visit • Board & staff conduct planning retreat • Design and implement funding strategy • Design and implement volunteer recruitment and training • Secure facility for clinic • Create an evaluation plan • Design and implement PR campaign	• # of patients referred from ER to the clinic/year • # of qualified patients enrolled in the clinic/year • # of patient visits/year • # of medical volunteers serving/year • # of patient flyers distributed • # of calls/month seeking info about clinic	• Memorandum of Agreement for free clinic space • Change in patient attitude about need for medical home • Change in # of scheduled annual physicals/follow-ups • Increased # of ER/physician referrals • Decreased volume of un-reimbursed emergencies treated in Memorial ER	• Patient co-payments supply 20% of clinic operating costs • 25% reduction in # of uninsured ER visits/year • 300 medical volunteers serving regularly each year • Clinic is a United Way Agency • Clinic endowment established • 90% patient satisfaction for 5 years. • 900 patients served/year

Chapter 3

Developing a Theory-of-Change Logic Model For Your Program

Drawing a picture of why your program should succeed

Whether you are a grantseeker developing a proposal for start-up funds or a grantee with a program already in operation, developing a logic model can strengthen your program. Logic models help identify the factors that will affect your program and enable you to anticipate the data and resources you will need to achieve success. As you engage in the process of creating your program logic model, your organization will systematically address these important program planning and evaluation issues:

- Description of the change strategy that your program supports.
- Definition of the problem you are attempting to address.
- Quantification of the scope of the needs or assets that make the case for your selection of the problem you address.
- Acknowledgement of the factors that may influence your ability to create change in your community.
- Application of best practice research that supports plausible solution strategies for identified problem area.
- Statement of your assumptions about why your selected strategies will work in your community in the ways you described.

Exercise 3: Program Planning constructs a program theory. Successful programs create change and are built on a solid knowledge of what works – your program's theory. Exercise 3 guides you through a series of six steps that diagram the fundamental theory that supports your program. This supports and builds upon the basic logic model. In most cases, if you are developing a new program, this step should come first to inform your preliminary thinking. We have placed it after basic logic models because it is a slightly more complex exercise.

Chapter 3

Exercise 3 – Constructing a Program Theory

Program Planning

The Importance of Framing Your Problems or Issues with Sound Program Theory

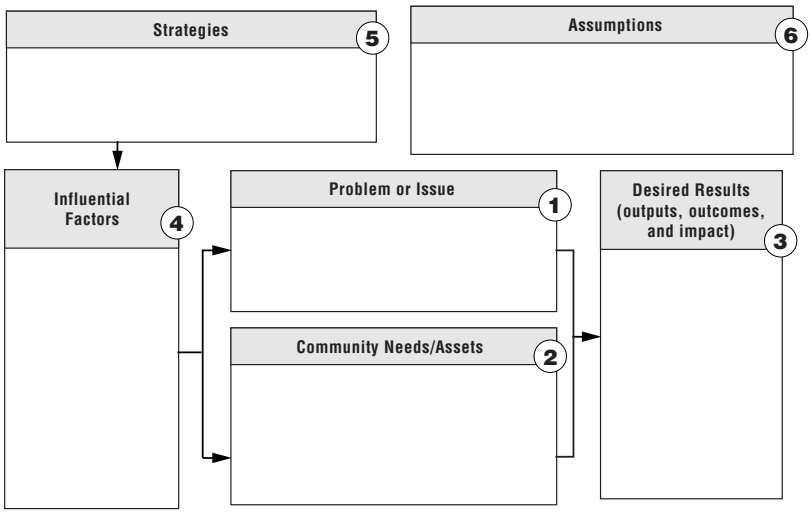
Imagine you work for a funding organization. Each quarter you review a mountain of grant proposals from worthy organizations doing good work. All of them need money. What information would you require to determine which programs to fund? Funders tell us they look for organizations that have done their homework by clearly defining the problem they plan to address, describing the reasons behind their approach, and outlining how they anticipate measuring their achievements. Funding decisions are more favorable if you can demonstrate clearly how and why they will succeed. Logic models help you do just that.

Most grantees know what they want to do in their communities; Exercise 1 makes a sound case for *how* and *why* funders should invest in your program.

It is crucial to begin program design with the basics. Funders encourage grantees to start by clearly and succinctly explaining the problems they plan to address. Completing Exercise 1 describes the issues your program will address, identifies the needs and assets of your community that are related to your issues, and specifies why certain results are desired. Funders and donors generally limit their investments to certain areas of interest, so if your program addresses several issues, you may want to construct a logic model for each one.

Exercise 3 Uses The Theory-of-Change Template

Logic Model Development
Program Planning Template – Exercise 1



INSTRUCTIONS: Exercise 3 will use the Theory-of-Change Template. In particular, you will use the information presented in the gray text boxes that follow about the Mytown example program to determine what theory-of-change was used to design and develop this program. Example information about influential factors, the problem, community needs/assets, strategies, and assumptions are provided. You will fill in the blank Theory-of-Change Template provided in the Forms Index (p. 57) to illustrate the program theory for the Mytown example. You can then look at the completed template on page 34 to compare your interpretation with that produced by the Mytown folks.

What problems are you attempting to solve or what issues are you striving to address? A well-constructed program theory points toward your program's eventual effectiveness. Begin your problem statement explaining concisely the issue you will address, stating the issue either as a community problem or asset. Your theory-of-change logic model will be built upon this statement, which illustrates how the program will function and what it expects to achieve in your community. It is smart to refer to research about your program's problem or issue in your statement; Internet searches can provide other successful program or "best practice" information.

PROBLEM STATEMENT EXAMPLE: There are increasing numbers of uninsured male workers, aged 40–55, in Mytown, USA, due to local plant closings. As the bottom line of hospitals shrink, the costs of uninsured care in local emergency rooms are negatively affecting local health systems. To meet the human and financial needs of Mytown, USA, an accessible, free medical home must be created to offer medical care and health education for Mytown's uninsured residents.

Insert Mytown's Problem or Issue in the Problem or Issue box of the Theory-of-Change Template

What needs or assets led you to address this issue? If a community needs assessment has been conducted or if you have prioritized community needs and capacity, data exist that make your case stronger and more specific by identifying and targeting your program's participants and activities. Documentation of community needs and assets also helps your evaluation plan later on. It can become a baseline providing indicators that measure progress made by your program over time. (Discussed in more detail in Chapter 4.)

DOCUMENTED NEEDS/ASSETS EXAMPLE: Memorial Hospital's Annual Report states that 28% of uninsured male patients, aged 40–55, received emergency room care in the previous year. Last year's United Way Community Needs Assessment identified health care for the uninsured as the #1 community health care issue. The Medical Society and Memorial Hospital's Task Force on the Uninsured is researching ways to address the needs of the uninsured AND reduce costly, inappropriate ER use.

Insert Mytown's community needs/assets in the Community Needs/Assets box of the Theory-of-Change Template.

What are your desired results? Identify what you expect your program to achieve in the near and longer term. These become your outputs, outcomes and impact.

DESIRED RESULTS EXAMPLE: Increase accessible, affordable health care for the uninsured and reduce the incidence of un-reimbursed care provided in emergency rooms. Create a free clinic that combines an appropriate, accessible, free medical home and patient education to reduce the numbers of uninsured males, aged 40–55, seeking care in emergency rooms. Anticipate a 15% *in*crease in males, aged 40–55, with a free medical home and a 25% *de*crease in the incidence of uninsured men seeking care in the ER within 5 years.

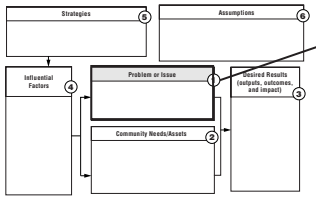
Insert Mytown's desired results (notice these are targeting men, which is more specific than in the basic logic model example) in Desired Results box of the Theory-of-Change Template.

What influential factors (protective and/or risk) could influence change in your community? What are the potential barriers and/or supports that might impact the change you hope for? Are there policies or other factors that could affect your program?

INFLUENTIAL FACTOR EXAMPLE: There is documented need for a free clinic. In its *Report for the New Millennium*, the Mytown Chamber of Commerce projects a 35% increase in the number of small businesses unable to afford employee health care benefits over the next five years. There is strong community support for a free clinic. At the request of Mytown United Way, Memorial Hospital and The Medical Society have created a joint task force to explore the creation of a free clinic.

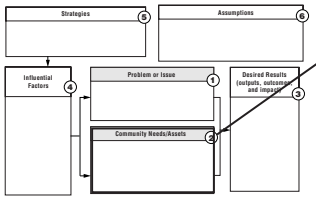
Insert Mytown's influential factors in the Influential Factors box of the Theory-of-Change Template.

Program Planning – Clarifying Program Theory



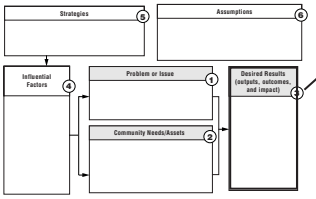
1 Problem or Issue Statement

Describe the problem(s) your program is attempting to solve or the issue(s) your program will address.



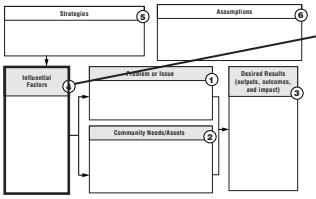
2 Community Needs/Assets

Specify the needs and/or assets of your community that led your program to address the problem(s) or issue(s).



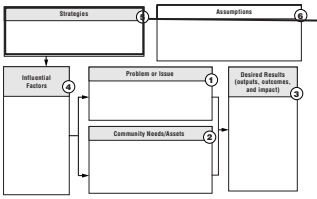
3 Desired Results (Outputs, Outcomes, and Impacts)

Identify your desired results, or vision of the future, by describing what you expect to achieve, near- or long-term, if your program is funded.



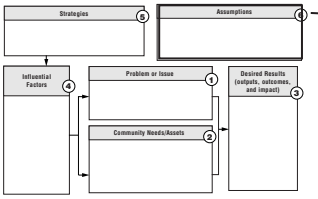
4 Influential Factors

List the factors (e.g., protective or risk factors, existing policy environment, or other factors) you believe will influence change in your community.



5 Strategies

List general, successful strategies or "best practices" your research identified that have helped communities like yours achieve the kinds of results your program promises.



6 Assumptions

State the assumptions behind *how* and *why* the identified change strategies will work in your community (e.g., principles, beliefs, ideas).

For more detail, see the Program Planning Template – Exercise 3 on p. 34.

Flowchart for Exercise 3

Why do you believe your program will work? Look for strong rationale based on "best practice" research that connects what you plan to do with *why* your approach will succeed. Funders are eager for evidence that supports why you propose the solutions you do. It's a good idea to relate your approach to similar change strategies that have proven effective in communities like yours. Reviewing literature and past evaluation reports from other programs (or your own work) will provide you with ample information to construct your program rationale. The Internet makes it easier to research effective program strategies.

PROGRAM STRATEGY EXAMPLE: A clinic using volunteer medical professionals reduced emergency room care visits in Anothertown, USA, in 1997 by 25%. A free clinic in Mytown, USA, using volunteer medical professionals could provide crucial, affordable medical homes for growing numbers of uninsured residents preventing costly, inappropriate emergency room use by males 40–55 experiencing coronary emergencies.

Insert Mytown's strategies in the Strategies box of the Theory-of-Change Template.

Why will your approach be effective? After you make the case for selecting a specific strategy from among the alternatives you researched, state out loud why this strategy is needed and why it will work in your community. It is important early on to document instances that describe the general condition of public reaction to your problem/issue and possible solutions.

You should draw direct conclusions about the statement of need and capacities in your community in your assumption. In addition, it should be quite apparent how your program intends to function as an intervention – to solve identified problems or build existing assets.

We list assumptions last in this exercise because in this abstracted learning format, the logic modeler has the benefit of all the information that supports assumptions. They are easier to spot and articulate with all the facts in front of you. In real-world conditions, assumption are best stated up-front – much earlier in the logic model development process – many basic logic models we have seen include a supporting page with the diagram that lists the assumptions that belie the model drawn.

ASSUMPTION EXAMPLE: As proven in Anothertown, access to affordable medical care reduces the incidence of emergency visits by providing appropriate, preventive primary care. A free medical clinic should prove successful in Mytown, because of its history of extraordinary volunteerism. Mytown's Medical Society officially encourages its 400 medical professional members to volunteer 20 hours each year to help medically underserved community residents. Mytown's Nursing Association is also interested in collaborating with a free clinic. Memorial Hospital has agreed to assist in planning and funding a free clinic. There is precedence for lending free facilities to medical projects serving those in need. Mytown's technical college donates space for Mytown's volunteer dental clinic. Mytown's Free Clinic will be strongly supported by the people, businesses and institutions of Mytown, USA.

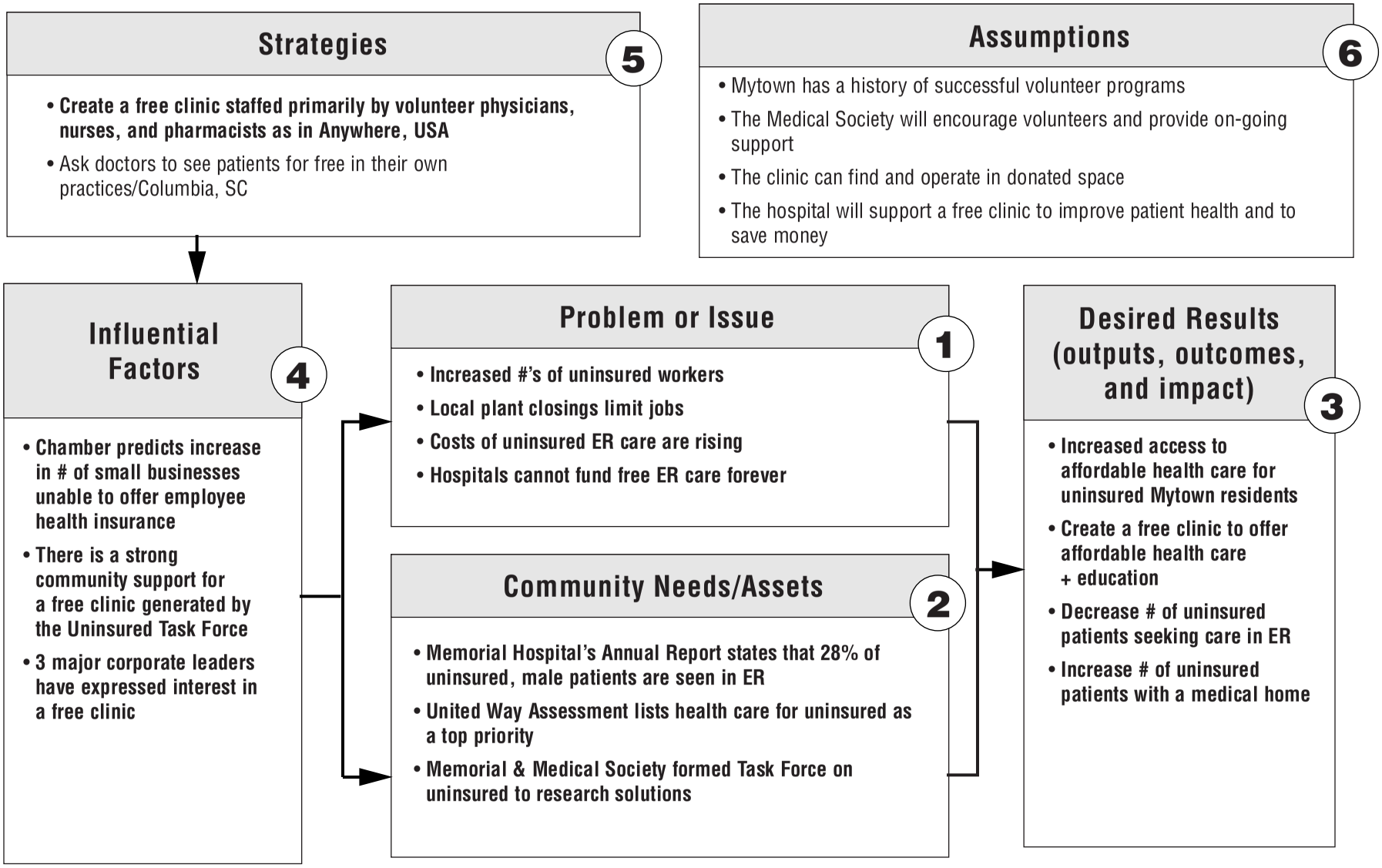
Insert Mytown's assumptions in the Assumptions box of the Theory-of-Change Template.

Exercise 3 Checklist:

After completing Exercise 3 and constructing your program theory, you can use the following checklist to assess the quality of your draft. It's helpful if someone other than the model's creators reviews the first program draft and completes the checklist, too.

	Exercise Three Checklist	Yes	Not Yet	Comments Revisions
1.	The problems to be solved/or issues to be addressed by the planned program are clearly stated	☐	☐	
2.	There is a specific, clear connection between the identified community needs/assets and the problems to be solved (or issues to be addressed).	☐	☐	
3.	The breadth of community needs/assets has been identified by expert/practitioner wisdom, a needs assessment and/or asset mapping process.	☐	☐	
4.	The desired results/changes in the community and/or vision for the future ultimately sought by program developers are specific.	☐	☐	
5.	Influential factors have been identified and cited from expert/practitioner wisdom and/or a literature review.	☐	☐	
6.	Change strategies are identified and cited from expert/practitioner wisdom and/or literature review.	☐	☐	
7.	The connection among known influential factors and broad change strategies has been identified.	☐	☐	
8.	The assumptions held for how and why identified change strategies should work in the community are clear.	☐	☐	
9.	There is consensus among stakeholders that the model accurately describes the proposed program and its intended results.	☐	☐	

Logic Model Development
Program Planning Template – Exercise 3



Chapter 4

Using Your Logic Model to Plan for Evaluation

Thinking through program evaluation questions in terms of the logic model components you have developed can provide the framework for your evaluation plan. Having a framework increases your evaluation's effectiveness by focusing on questions that have real value for your stakeholders.

- Prioritization of where investment in evaluation activities will contribute the most useful information for program stakeholders.
- Description of your approach to evaluation.

There are two exercises in this chapter; Exercise 4 deals with posing evaluation questions and Exercise 5 examines the selection of indicators of progress that link back to the basic logic model or the theory-of-change model depending on the focus of the evaluation and its intended primary audiences.

Exercise 4 – Posing Evaluation Questions

The Importance of "Prove" and "Improve" Questions

There are two different types of evaluation questions – ***formative*** help you to *improve* your program and ***summative*** help you *prove* whether your program worked the way you planned. Both kinds of evaluation questions generate information that determines the extent to which your program has had the success you expected and provide a groundwork for sharing with others the successes and lessons learned from your program.

Benefits of Formative and Summative Evaluation Questions [3]

Formative Evaluation – Improve	Summative Evaluation – Prove
Provides information that helps you improve your program. Generates periodic reports. Information can be shared quickly.	Generates information that can be used to demonstrate the results of your program to funders and your community.
Focuses most on program activities, outputs, and short-term outcomes for the purpose of monitoring progress and making mid-course corrections when needed.	Focuses most on program's intermediate-term outcomes and impact. Although data may be collected throughout the program, the purpose is to determine the value and worth of a program based on results.
Helpful in bringing suggestions for improvement to the attention of staff.	Helpful in describing the quality and effectiveness of your program by documenting its impact on participants and the community.

[3] Adapted from Bond, S.L., Boyd, S. E., & Montgomery, D.L.(1997 *Taking Stock: A Practical Guide to Evaluating Your Own Programs, Chapel Hill, NC:* Horizon Research, Inc. Available online at http://www.horizon-research.com.

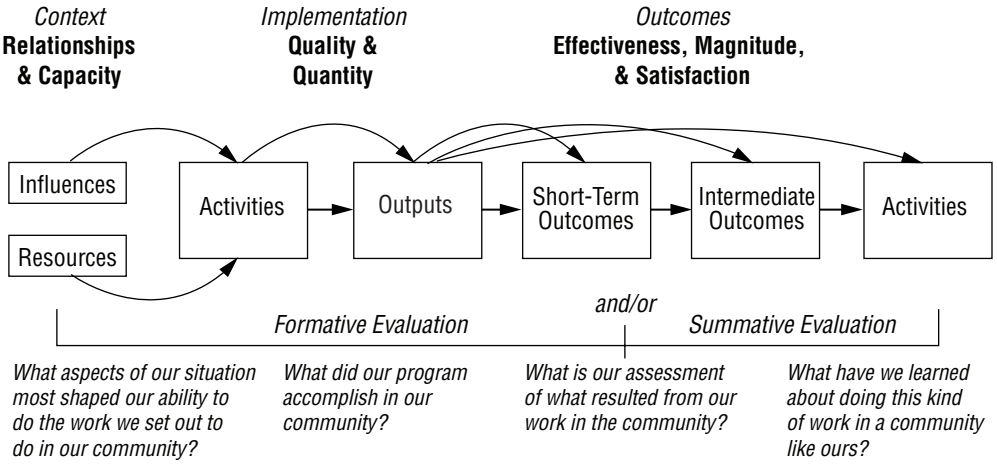
Chapter 4

Looking at Evaluation from Various Vantage Points

How will you measure your success? What will those investing in your program or your target audience want to know?

A clear logic model illustrates the purpose and content of your program and makes it easier to develop meaningful evaluation questions from a variety of program vantage points: context, implementation and results (which includes outputs, outcomes, and impact).

What Parts of Your Program Will Be Evaluated?
Using a logic model to frame your evaluation questions.



Remember you can draw upon the basic logic model in Exercises 1 and 2 and the theory-of-change model in Exercise 3. Feasibility studies and needs assessments serve as valuable resources for baseline information on influences and resources collected during program planning.

Context **is how the program functions within the economic, social, and political environment of its community** and addresses questions that explore issues of program relationships and capacity. What factors might influence your ability to do the work you have planned? These kinds of evaluation questions can help you explain some of the strengths and weakness of your program as well as the effect of unanticipated and external influences on it.

Sample CONTEXT QUESTIONS: Can we secure a donated facility? With the low morale created by high unemployment, can we secure the financial and volunteer support we need? How many medical volunteers can we recruit? How many will be needed each evening? How will potential patients find out about the clinic? What kind of medical care will patients need? How can we let possible referral sources know about the clinic and its services? What supplies will we need and how will we solicit suppliers for them? What is it about the free clinic that supports its ability to reduce the numbers of patients seeking care in Memorial Hospital's ER?

Implementation **assesses the extent to which activities were executed as planned**, since a program's ability to deliver its desired results depends on whether activities result in the quality and quantity of outputs specified. They tell the story of your program in terms of what happened and why.

SAMPLE IMPLEMENTATION QUESTIONS: What facility was secured? How many patients were seen each night/month/year? What organizations most frequently referred patients to the clinic? How did patients find out about the clinic? How many medical volunteers serve each night/month/year? What was the value of their services? What was the most common diagnosis? What supplies were donated? How many patients per year did the Clinic see in its first/second/third year?

Outcomes **determine the extent to which progress is being made toward the desired changes in individuals, organizations, communities, or systems.** Outcome questions seek to document the changes that occur in your community as a result of your program. Usually these questions generate answers about effectiveness of activities in producing changes in magnitude or satisfaction with changes related to the issues central to your program.

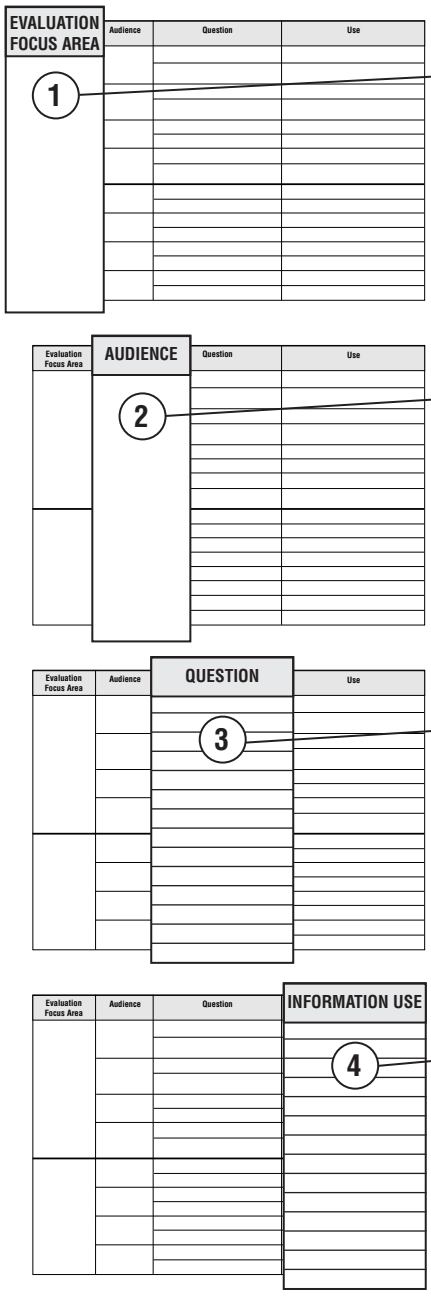
SAMPLE OUTCOME QUESTIONS: How many inappropriate, uninsured patients sought medical care in Memorial's ER in the Clinic's first/second/third year? Was there a reduction in unfunded ER visits? How did the number of uninsured patients compare to previous years when the clinic was not operating? What was the cost/visit in the Free Clinic? What is the cost/visit in Memorial's ER? How do they compare? What were the cost savings to Memorial Hospital? How satisfied were Clinic patients with the care they received? How satisfied were volunteers with their service to the Clinic?

Chapter 4

Creating Focus

Though it is rare, you may find that examining certain components of your program is sufficient to satisfy your information needs. Most often, however, you will systematically develop a series of evaluation questions, as shown in the Flowchart for Evaluation Question Development.

Flowchart for Evaluation Question Development



1 Evaluation Focus Area

What is going to be evaluated? List those components from your theory and/or logic model that you think are the most important aspects of your program. These areas will become the focus of your evaluation.

2 Audience

What key audience will have questions about your focus areas? For each focus area you have identified, list the audiences that are likely to be the most interested in that area.

3 Question

What questions will your key audience have about your program? For each focus area and audience that you have identified, list the questions they might have about your program.

4 Information Use

If you answer a given question, what will that information be used for? For each audience and question you have identified, list the ways and extent to which you plan to make use of the evaluation information.

The use of program theory as a map for evaluation doesn't necessarily imply that every step of every possible theory has to be studied. ...Choices have to be made in designing an evaluation about which lines of inquiry to pursue. ...The theory provides a picture of the whole intellectual landscape so that people can make choices with a full awareness of what they are ignoring as well as what they are choosing to study... Weiss (1998)

Evaluation

For more detail, see the Evaluation Planning Template – Exercise 4 on p. 44.

What is going to be evaluated?

For each area on which your program focuses, list the most important aspects of your program theory and logic model. Focus your evaluation on them.

Focus Area Examples:

Context Examples — Evaluating relationships and capacity. How will the Free Clinic recruit and train effective board and staff members? What is the best way to recruit, manage, retain and recognize medical and administrative volunteers and other Clinic partners? What is the most effective way to recruit and retain uninsured patients? How will the operation of a Free Clinic affect Memorial Hospital's expenses for providing uninsured medical care in its ER? How many patients can Clinic volunteers effectively serve on a regular basis? What is the ideal patient/volunteer ratio?

Insert focus areas into Focus Area Column of Evaluation Questions Development Template for Evaluation Planning, Exercise 4.

Implementation Examples – Assessing quality and quantity. How many major funding partners does the clinic have? How are volunteers and patients scheduled? How many medical volunteers serve Clinic patients on a regular basis? What is the value of their services? What is the most common diagnosis at the Clinic? What is the most common diagnosis of uninsured patients seen in Memorial's ER? How long do patients wait to be seen at the Clinic? Is there a patient or volunteer waiting list?

Insert focus areas into Focus Area Column of Evaluation Questions Development Template for Evaluation Planning, Exercise 4.

Outcomes – Measuring effectiveness, magnitude and satisfaction. Has the clinic increased access to care for a significant number of Mytown's uninsured citizens? How many residents of Mytown, USA, do not have health insurance? How many patients does the Clinic serve on a regular basis? What is that ratio? What is the cost per visit in the Clinic and Memorial's ER? How do the costs compare? What is the satisfaction level of Clinic patients and volunteers with Clinic services and facilities? How many donors does the Clinic have? What is their satisfaction with Clinic services and facilities? How effectively is the Clinic educating, engaging and involving its partners? What organizations have officially endorsed the Clinic? What is the board and staff's satisfaction with clinic operations, facilities and services?

Insert focus areas into Focus Area Column of Evaluation Questions Development Template for Evaluation Planning, Exercise 4.

The benefits of asking and answering evaluation questions depend on how clear you are about the purpose of your evaluation, who needs to know what when, and the resources you have available to support the evaluation process.

Chapter 4

What Information Will Your Program's Audiences Want?

As shown below, program audiences will be interested in a variety of different kinds of information. Donors may want to know if their money did what you promised it would. Patients might want to know how many patients the clinic serves and how many volunteers it has. Physicians donating their time and talent could be interested in the financial value of their contributions. If you ask your audiences what they want to know, you'll be sure to build in ways to gather the evaluation data required.

Audience	Typical Questions	Evaluation Use
Program Management and Staff	Are we reaching our target population? Are our participants satisfied with our program? Is the program being run efficiently? How can we improve our program?	Programming decisions, day-to-day operations
Participants	Programming decisions, day-to-day operations Did the program help me and people like me? What would improve the program next time?	Decisions about continuing participation.
Community Members	Is the program suited to our community needs? What is the program really accomplishing?	Decisions about participation and support.
Public Officials	Who is the program serving? What difference has the program made? Is the program reaching its target population? What do participants think about the program? Is the program worth the cost?	Decisions about commitment and support. Knowledge about the utility and feasibility of the program approach.
Funders	Is what was promised being achieved? Is the program working? Is the program worth the cost?	Accountability and improvement of future grantmaking efforts.

How often do you have to gather data? Whether a question is more formative or summative in nature offers a clue on when information should be collected.

- Formative information should be periodic and reported/shared quickly to improve your efforts.
- Summative tends to be "before and after" snapshots reported after the conclusion of the program to document the effectiveness and lessons learned from your experience.

Involve Your Audience in Setting Priorities

Program developers often interview program funders, participants, staff, board and partners to brainstorm a list of all possible questions for a key area identified from their program theory or from their logic models. That list helps determine the focus of the evaluation. Involving your audience from the beginning makes sure you gather meaningful information in which your supporters have a real interest.

Prioritization is a critical step. No evaluation can answer all of the questions your program's audiences may ask. The following questions can help you narrow your number of indicators: How many audiences are interested in this information? Could knowing the answer to this question improve your program? Will this information assess your program's effectiveness?

The final focus for your evaluation is often negotiated among stakeholders. It is important to keep your evaluation manageable. **It is preferable to answer a few important questions thoroughly than to answer several questions poorly**. How well you can answer your questions will depend on the time, money, and expertise you have at your disposal to perform the functions required by the evaluation.

What key audiences will have questions about your evaluation focus areas?

For each focus area that you identified in the previous step, list the audiences that are likely to be most interested in that area. Summarize your audiences and transfer to the Audience Column of the Evaluation Questions Development Template for Evaluation Planning, Exercise 4.

Context – Relationships and Capacity

Example audiences: Medical professionals, Memorial Hospital Board and Staff (especially ER staff), Medical associations, Foundations, The Chamber of Commerce, United Way, The Technical College, uninsured residents, medical supply companies, local media, public officials.

Implementation – Quality and Quantity

Example audiences: Funders, In-kind donors, Medical and administrative volunteers, Board, Staff, Patients, Public Officials, The media, Medical associations, Local businesses, Health care organizations.

Outcomes – Effectiveness, Magnitude, and Satisfaction

Example audiences: Funders, In-kind donors, Volunteers, Board, Staff, Patients, Public Officials, The media, Medical associations, Local businesses, Health care organizations.

What questions will key audiences ask about your program?

For each focus area and key audience you identified in the previous step, list the questions your stakeholders ask about your program. Insert summaries in the Question Column of the Evaluation Questions Development Template for Evaluation Planning, Exercise 4 (on page 44).

Sample of Key Audience Questions:

- Who are the collaborative partners for this program? What do they provide?
- What is the budget for this program?
- How many staff members does the program have?
- How many patients does the clinic serve?
- How many visits per year does the average patient have?
- What is the most common diagnosis?
- Does the clinic save the hospital money?
- How does the organization undertake and support program evaluation?
- How are medical volunteers protected from lawsuits?
- How satisfied are patients, volunteers, board and staff with the clinic's services?
- What do experts say about the clinic?
- How many uninsured patients still seek inappropriate care in the ER? Why?

How will the evaluation's information be used?

For each question and audience you identified in the previous step, list the ways and extent to which you plan to make use of the evaluation information. Summarize audience use of information. Insert in the Use Column of the Evaluation Questions Development Template for Evaluation Planning, Exercise 4.

Context – Relationships and Capacity Examples

- Measure the level of community support.
- Assess effectiveness of community outreach.
- Assess sustainability of Clinic funding sources.
- Improve volunteer and patient recruitment methods.
- Secure additional Clinic partners.

Implementation – Quality and Quantity Examples

- Assess optimal number of volunteers and patients to schedule per session to improve operating effectiveness while maintaining patient and volunteer satisfaction.
- Measure patient, volunteer, staff, board, donors and community satisfaction with clinic.
- Determine cost savings per visit. Share information with local medical and business groups to encourage their support.

Outcomes and Impact – Examples of Effectiveness, Magnitude, and Satisfaction

- Cost savings of Clinic – use to obtain additional volunteer and financial support from Memorial Hospital.
- Patient satisfaction survey results – use to improve patient services and satisfaction.
- Analysis of most frequent referral sources – use to present information seminars to ER staff, social service workers and unemployment insurance clerks to increase patient referrals and intakes.
- Analysis of most prevalent patient diagnoses – use to create relevant patient health education newsletter. Patient tracking system will measure impact of education program.

Exercise 4 Checklist: After completing Exercise 4 you can use the following checklist to assess the quality of your draft.

Posing Questions Quality Criteria	Yes	Not Yet	Comments Revisions
1. A variety of audiences are taken into consideration when specifying questions.	☐	☐	
2. Questions selected are those with the highest priority.	☐	☐	
3. Each question chosen gathers useful information.	☐	☐	
4. Each question asks only one question (i.e. "extent of X, Y, and Z" is not appropriate).	☐	☐	
5. It is clear how the question relates to the program's logic model.	☐	☐	
6. The questions are specific about what information is needed.	☐	☐	
7. Questions capture lessons learned about your work along the way.	☐	☐	
8. Questions capture lessons learned about your program theory along the way.	☐	☐	

Logic Model Development
Evaluation Planning Template – Exercise 4

Evaluation Focus Area	Audience	Question	Use
Relationships	Funders	Is the program cost effective?	Cost benefits/fundraising
		Are volunteers & patients satisfied with Clinic services?	Program promotion/fundraising
	Medical Volunteers	What is the most common diagnosis?	Quality assurance/Planning
		How will medical volunteers be protected from lawsuits?	Volunteer recruitment
	Patients	Am I receiving quality care?	Program improvement & planning
		How long can I receive care here?	" "
	Staff	Are we reaching our target population?	Evaluation/program promotion
		How do patients find us? What's our best promotional approach?	Evaluation and/or improvement
Outcomes	Funders/Donors	Program Budget?	Cost benefit analysis
		Cost/visit?	"
	Volunteers	Visits/month/year?	Annual Report/program promotion/public relations
		Cost savings for Memorial Hospital?	Annual Report/program promotion/fundraising
	Patients	Volunteers/year?	Annual Report/volunteer recruitment
		Patient satisfaction	Program improvements/staff training
	Staff	Patient & volunteer satisfaction	"
		Common DRG(?)	

Exercise 5 – Establishing Indicators

One of the biggest challenges in developing an evaluation plan is choosing what kind of information best answers the questions you have posed. **It is important to have general agreement across your audiences on what success will look like. Indicators** are the measures you select as markers of your success.

The biggest problem is usually that people are trying to accomplish too many results. Once they engage in a discussion of indicators, they start to realize how much more clarity they need in their activities.

I also find that it is important that the program, not the evaluator, is identifying the indicators. Otherwise, the program can easily discredit the evaluation by saying they don't think the indicators are important, valid, etc.

Beverly Anderson Parsons, WKKF Cluster Evaluator

In this last exercise you create a set of indicators. They are often used as the starting point for designing the data collection and reporting strategies (e.g., the number of uninsured adults nationally, statewide, in Mytown, USA, or the number of licensed physicians in Mytown). Often organizations hire consultants or seek guidance from local experts to conduct their evaluations. Whether or not you want help will depend on your organization's level of comfort with evaluation and the evaluation expertise among your staff.

Focus Area	Indicators	How to Evaluate[1]
Influential Factors	Measures of influential factors – may require general population surveys and/or comparison with national data sets[2].	Compare the nature and extent of influences before (baseline) and after the program.
Resources	Logs or reports of financial/staffing status.	Compare actual resources acquired against anticipated.
Activities	Descriptions of planned activities. Logs or reports of actual activities. Descriptions of participants.	Compare actual activities provided, types of participants reached against what was proposed.
Outputs	Logs or reports of actual activities. Actual products delivered.	Compare the quality and quantity of actual delivery against expected.
Outcomes & Impacts	Participant attitudes, knowledge, skills, intentions, and/or behaviors thought to result from your activities[3].	Compare the measures before and after the program[4].

Examples and Use of Indicators.

[1] This table was adapted from *A Hands-on Guide to Planning and Evaluation* (1993) available from the National AIDS Clearinghouse, Canada.
[2] You may want to allocate resources to allow for the assistance of an external evaluation consultant to access national databases or perform statistical analyses.
[3] Many types of outcomes and impact instruments (i.e. reliable and valid surveys and questionnaires) are readily available. The Mental Measurement Yearbook published by the Buros Institute (http://www.unl.edu/buros/) and the ERIC Clearinghouse on Assessment and Evaluation (http://ericae.net/) are great places to start.
[4] You may need to allocate resources to allow for the assistance of an external evaluation consultant.

Chapter 4

Our advice is to keep your evaluation simple and straightforward. The logic model techniques you have been practicing will take you a long way toward developing an evaluation plan that is meaningful *and* manageable.

Determine the kinds of data you will need and design methods to gather the data (i.e., patient registration forms, volunteer registration forms, daily sign-in sheets, national, state and local statistics). Sometimes, once an indicator (type of data) is selected, program planners set a specific target to be reached as an agreed upon measure of success (for example 25% decrease in the numbers of inappropriate ER visits).

As in the previous exercises use the space below to loosely organize your thoughts. Then, once the exercise is completed and assessed, use the **Indicator Development Template** on page 61 to record your indicators and technical assistance needs.

Filling in the Flowchart for Indicator Development

What information will be gathered to "indicate" the status of your program and/or its participants?

Focus Area	Question	Indicators	Technical Assistance Needed
1	2	3	4

Column 1: *Focus Areas* – From the information gathered in Exercise 4, transfer the areas on which your evaluation will focus into column one (for example, patient health, volunteer participation, sustaining supporting partnerships).

Column 2: *Questions* – transfer from Exercise 4 the major questions related to each focus area – big questions your key audiences want answered. Remember to keep your evaluation as simple as possible.

Column 3: *Indicators* – Specify the indicators (types of data) against which you will measure the success/progress of your program. It's often helpful to record the sources of data you plan to use as indicators (where you are likely to find or get access to these data).

Column 4: *Technical Assistance* – To what extent does your organization have the evaluation and data management expertise needed to collect and analyze the data that related to each indicator? List any assistance that would be helpful – universities, consultants, national and state data experts, foundation evaluation departments, etc.

Exercise 5 Checklist: Review what you have created using the checklist below to assess the quality of your evaluation plan.

Establishing Indicators Quality Criteria		Yes	Not Yet	Comments Revisions
1.	The focus areas reflect the questions asked by a variety of audiences. Indicators respond to the identified focus areas and questions.	☐	☐	
2.	Indicators are SMART–Specific, Measurable, Action-oriented, Realistic, and Timed.	☐	☐	
3.	The cost of collecting data on the indicators is within the evaluation budget.	☐	☐	
4.	Source of data is known.	☐	☐	
5.	It is clear what data collection, management, and analysis strategies will be most appropriate for each indicator.	☐	☐	
6.	Strategies and required technical assistance have been identified and are within the evaluation budget for the program.	☐	☐	
7.	The technical assistance needed is available.	☐	☐	

Logic Model Development
Indicators Development Template – Exercise 5

Focus Area	Question	Indicators	Technical Assistance Needed
Relationships	Are volunteers & patients satisfied w/ clinic care?	• Patient satisfaction surveys • Volunteer satisfaction tests	Anywhere's patient satisfaction surveys Anywhere's volunteer survey
Outcomes	Are we reaching our target population?	• % of clinic patients vs. % of uninsured citizens in Mytown, USA • # of qualified clinic patients/year	Reports from Chamber of Commerce Patient database creation
	How do patients find the clinic?	• Annual analysis of telephone referral log • Referral question on patient intake form	Telephone log database Anywhere's patient intake form
	Does the clinic save the community $?	• Cost/visit • # of uninsured patients seen in hospital ER – beginning the year before clinic opened	Budget figures; patient service records Tracking database software Strategic direction for analysis
	What does the clinic provide?	• Most common diagnosis • Hospital cost/visit for common diagnosis	DRG workbook/tables (hospital staff) Input from hospital billing staff
	How has volunteering affected doctors, nurses, administrators and patients?	• Annual volunteer survey • Patient satisfaction survey • # of volunteers/year • # of volunteers donating to clinic operations	Anywhere surveys and analysis instruments Volunteer management database Donor database (Raiser's Edge?)

Resources Appendix

This Appendix provides information on print and electronic resources available to support you in your logic model development process.

1. Logic Model Information and Examples

University of Nevada, Reno Western CAPT web site
http://www.unr.edu/colleges/educ/captta/prev/evaluate.htm

BJA Evaluation web site
http://www.bja.evaluationwebsite.org

Schmitz, C. C. & Parsons, B. A. (1999). *Everything you wanted to know about logic models but were afraid to ask.* Battle Creek, MI: W.K. Kellogg Foundation.

2. United Way of America's Outcome Models

United Way of America web site
http://www.unitedway.org/outcomes/contents.htm

Measuring program outcomes: A practical approach.
United Way of America
701 North Fairfax Street
Alexandria, VA 22314
(703) 836-7100

3. Definitions and Information on Program Theory and Evaluation

Program Theory Definitions

- *A plausible and sensible model of how a program is supposed to work* (Bickman, 1987, p. 5).
- *The set of assumptions about the relationships between the strategy and tactics the program has adopted and the social benefits it is expected to produce* (Rossi, Freeman, & Lipsey, 1999, p.98).
- *The full chain of objectives that links inputs to activities, activities to ... outputs, ... outputs to ... outcomes, and ... outcomes to ultimate goals constitutes a program's theory* (Patton, 1997, p. 218).
- *A set of interrelated assumptions, principles, and/or propositions to explain or guide social actions* (Chen, 1990, p. 40).
- *An explanation of the causal links that tie program inputs to expected program outputs* (Weiss, 1998, p. 55).
- *A chain of causal assumptions linking program resources, activities, intermediate outcomes, and ultimate goals* (Wholey, 1987, p. 78).

Bickman, L. (Ed.). (1987). Using program theory in evaluation. *New Directions for Program Evaluation Series (no. 33)*. San Francisco: Jossey-Bass.

Chen, H. T. (1990). *Theory driven evaluations*. Newbury Park, CA: Sage.

Wholey, J. S. (Ed). (1987). *Organizational excellence: Stimulating quality and communicating value*. Lexington, MA: Lexington Books.

Wholey, J. S., Hatry, H. P., & Newcomer, K. E. (Eds.). (1994). *Handbook of Practical Program Evaluation*. San Francisco: Jossey-Bass Publishers.

Weiss, C. H. (1998). *Evaluation*: *Methods for studying programs and policies*. (2nd Ed). Upper Saddle River, NJ: Prentice Hall.

4. W.K. Kellogg Foundation Logic Model Examples

W.K. Kellogg Foundation Web site
http://www.wkkf.org
W.K. Kellogg Foundation Evaluation Handbook.

Parsons, B. A. (1999). *Making logic models more systemic*. A paper presented at the Annual Meeting of the American Evaluation Association, Orlando, FL, November 1999.

Parsons, B. A., Schmitz, Co (1999) *Everything You Wanted to Know About Logic Models But Were Afraid to Ask*. A paper presented at the Annual Meeting of the American Evaluation Association, Orlando, FL, November 1999.

5. Information about Logic Model Development and Use

The Evaluation Forum (1999). *Outcomes for success*.
The Evaluation Forum
1932 First Avenue, Suite 403
Seattle, WA 98101
(206) 269-0171

Freddolino, P. P. (1999). *The program logic model: What it is and how to teach it.* A preconference workshop presented at the 1999 Conference of the Michigan Association for Evaluation, East Lansing, MI, May.

Targeting Outcomes of Programs.
http://deal.unl.edu/TOP/synopsis.htm

Innovation Network, Inc. electronic logic model development tool
http://www.inetwork.org

6. Evaluation Planning Information

W.K. Kellogg Foundation Web site
http://www.wkkf.org
W.K. Kellogg Foundation Evaluation Handbook.

Taking Stock.
http://www.horizon-research.org

The Evaluation Forum (1994). *A field guide to outcome-based program evaluation*.

The Evaluation Forum
1932 First Avenue, Suite 403
Seattle, WA 98101
(206) 269-0171

Rossi, P. H., Freeman, H. E., & Lipsey, M. W. (1999). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage.

Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Thousand Oaks, CA: Sage.

Forms Appendix

This Appendix provides the worksheet templates and checklists for exercises 1-5:

Logic Model Development Program Planning and Implementation

Exercises 1 and 2 Template

Exercise 1 Checklist

Exercise 2 Checklist

Theory-of-Change Logic Model Development Planning

Exercise 3 Template

Exercise 3 Checklist

Logic Model Development Evaluation and Indicators Development

Exercise 4 Template

Exercise 4 Checklist

Exercise 5

Exercise 5 Checklist

Logic Model Development
Program Implementation Template – Exercise 1 & 2

RESOURCES	ACTIVITIES	OUTPUTS	SHORT- & LONG-TERM OUTCOMES	IMPACT
In order to accomplish our set of activities we will need the following:	*In order to address our problem or asset we will accomplish the following activities:*	*We expect that once accomplished these activities will produce the following evidence or service delivery:*	*We expect that if accomplished these activities will lead to the following changes in 1–3 then 4–6 years:*	*We expect that if accomplished these activities will lead to the following changes in 7–10 years:*

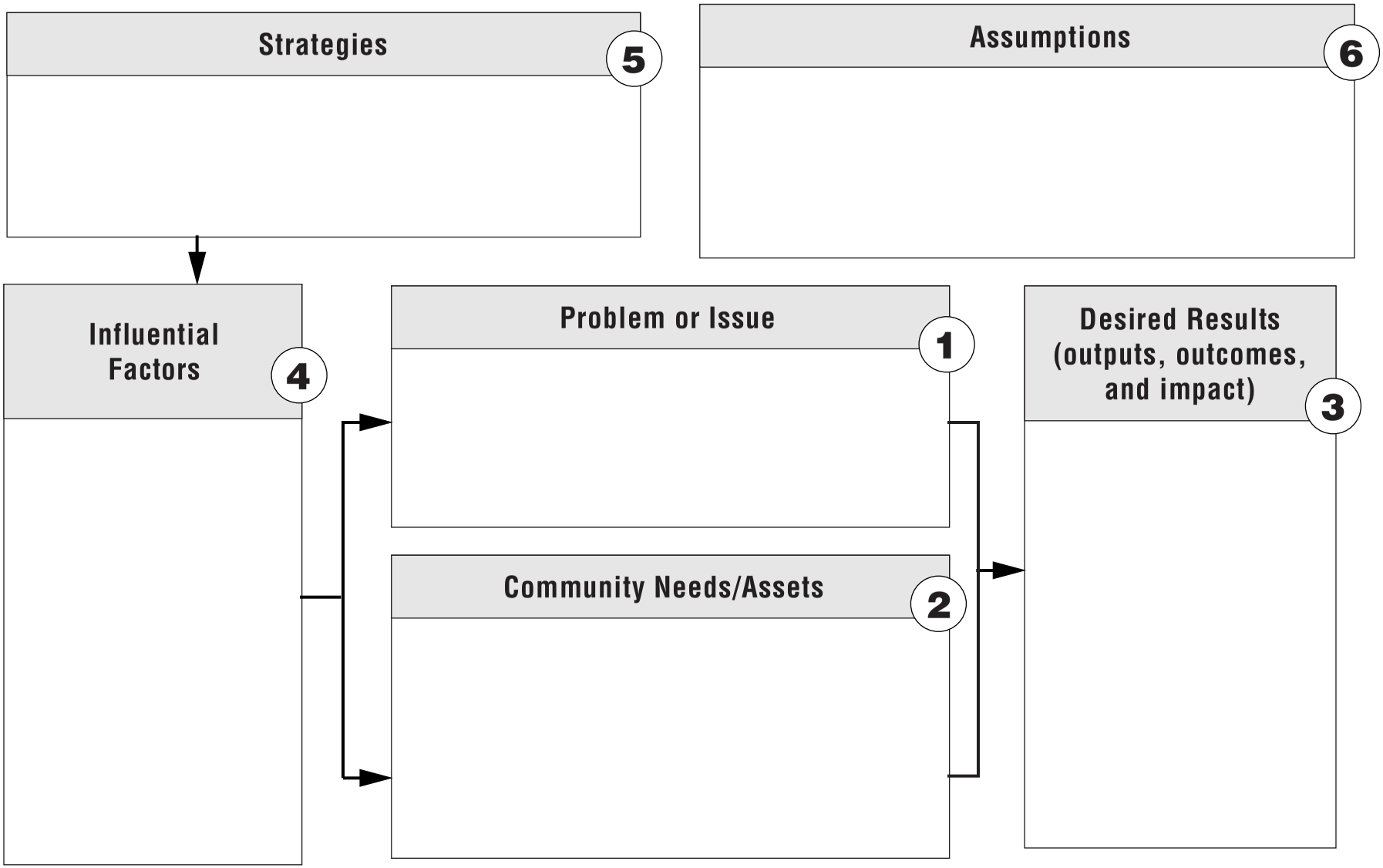
Exercise 1 Checklist

	Progress Toward Results Quality Criteria – 1	Yes	Not Yet	Comments/ Revisions
1.	A variety of audiences are taken into consideration when specifying credible outputs, outcomes, and impacts.	☐	☐	
2.	Target participants and/or partners are described and quantified as outputs (e.g. 100 teachers from 5 rural high schools).	☐	☐	
3.	Events, products, or services listed are described as outputs in terms of a treatment or dose (e.g. 30 farmers will participate in at least 3 sessions of program, or curriculum will be distributed to at least 12 agencies).	☐	☐	
4.	The intensity of the intervention or treatment is appropriate for the type of participant targeted (e.g. higher-risk participants warrant higher intensities).	☐	☐	
5.	The duration of the intervention or treatment is appropriate for the type of participant targeted (e.g. higher-risk participants warrant longer duration).	☐	☐	
6.	Outcomes reflect reasonable, progressive steps that participants can make toward longer-term results.	☐	☐	
7.	Outcomes address awareness, attitudes, perceptions, knowledge, skills, and/ or behavior of participants.	☐	☐	
8.	Outcomes are within the scope of the program's control or sphere of reasonable influence.	☐	☐	
9.	It seems fair or reasonable to hold the program accountable for the outcomes specified.	☐	☐	
10.	The outcomes are specific, measurable, action-oriented, realistic, and timed.	☐	☐	
11.	The outcomes are written as change statements (e.g. things increase, decrease, or stay the same).	☐	☐	
12.	The outcomes are achievable within the funding and reporting periods specified.	☐	☐	
13.	The impact, as specified, is not beyond the scope of the program to achieve.	☐	☐	

Exercise 2 Checklist

Theory into Action Quality Criteria	Yes	Not Yet	Comments/Revisions
1. Major activities needed to implement the program are listed.	☐	☐	
2. Activities are clearly connected to the specified program theory.	☐	☐	
3. Major resources needed to implement the program are listed.	☐	☐	
4. Resources match the type of program.			
5. All activities have sufficient and appropriate resources.	☐	☐	

Logic Model Development
Program Planning Template – Exercise 3



Exercise 3 Checklist

	Exercise Three Checklist	Yes	Not Yet	Comments/ Revisions
1.	The problems to be solved/or issues to be addressed by the planned program are clearly stated	☐	☐	
2.	There is a specific, clear connection between the identified community needs/assets and the problems to be solved (or issues to be addressed).	☐	☐	
3.	The breadth of community needs/assets has been identified by expert/practitioner wisdom, a needs assessment and/or asset mapping process.	☐	☐	
4.	The desired results/changes in the community and/or vision for the future ultimately sought by program developers are specific.	☐	☐	
5.	Influential factors have been identified and cited from expert/practitioner wisdom and/or a literature review.	☐	☐	
6.	Change strategies are identified and cited from expert/practitioner wisdom and/or literature review.	☐	☐	
7.	The connection among known influential factors and broad change strategies has been identified.	☐	☐	
8.	The assumptions held for how and why identified change strategies should work in the community are clear.	☐	☐	
9.	There is consensus among stakeholders that the model accurately describes the proposed program and its intended results.	☐	☐	

Logic Model Development
Evaluation Planning Template – Exercise 4

Evaluation Focus Area	Audience	Question	Use

Exercise 4 Checklist

Posing Questions Quality Criteria	Yes	Not Yet	Comments/ Revisions
1. A variety of audiences are taken into consideration when specifying questions.	☐	☐	
2. Questions selected are those with the highest priority.	☐	☐	
3. Each question chosen gathers useful information.	☐	☐	
4. Each question asks only one question (i.e. "extent of X, Y, and Z") is not appropriate).	☐	☐	
5. It is clear how the question relates to the program's logic model.	☐	☐	
6. The questions are specific about what information is needed.	☐	☐	
7. Questions capture lessons learned about your work along the way.	☐	☐	
8. Questions capture lessons learned about your program theory along the way.	☐	☐	

Logic Model Development
Indicators Development Template – Exercise 5

Focus Area	Question	Indicators	Technical Assistance Needed

Exercise 5 Checklist

Establishing Indicators Quality Criteria		Yes	Not Yet	Comments/ Revisions
1.	The focus areas reflect the questions asked by a variety of audiences. Indicators respond to the identified focus areas and questions.	☐	☐	
2.	Indicators are SMART–Specific, Measurable, Action-oriented, Realistic, and Timed.	☐	☐	
3.	The cost of collecting data on the indicators is within the evaluation budget.	☐	☐	
4.	Source of data is known.	☐	☐	
5.	It is clear what data collection, management, and analysis strategies will be most appropriate for each indicator.	☐	☐	
6.	Strategies and required technical assistance have been identified and are within the evaluation budget for the program.	☐	☐	
7.	The technical assistance needed is available.	☐	☐	

W.K.KELLOGG
FOUNDATION

One Michigan
Avenue East
Battle Creek, MI
49017-4012
USA
269-968-1611
TDD on site
Facsimile: 269-968-0413
www.wkkf.org

EV3815
Item #1209
0305 5M SCG
Printed on Recycled Paper

THE ETHELYN R. STRONG
SCHOOL OF SOCIAL WORK

Powerful Note-Taking

Skeletal Outline

Use this outline to take notes for the module. Each of the resources in the module is designed to support your learning of one or more of the learning objectives. Refer to the module overview and write down the learning objectives below. Keep these learning objectives in mind as you engage with the module activities and assignments.

Learning Objectives

1. _____
2. _____
3. _____

Reading (Title, Author, Date): _____

In your own words, explain what you learned in relation to the learning objective(s).

What are you still confused about?

Create a question or set of questions you think the instructor might ask at the end of this reading and include what you believe to be an appropriate response or set of responses.

Module Resources

Use the space below to keep track of any additional resources provided in the module that help you better understand the learning objectives for this module. Specify where to find the resource and the information you find useful so you can refer to it later.

Resource title/location:

_____

Useful information in this resource:

_____

Resource title/location:

_____

Useful information in this resource:

_____

Resource title/location:

_____

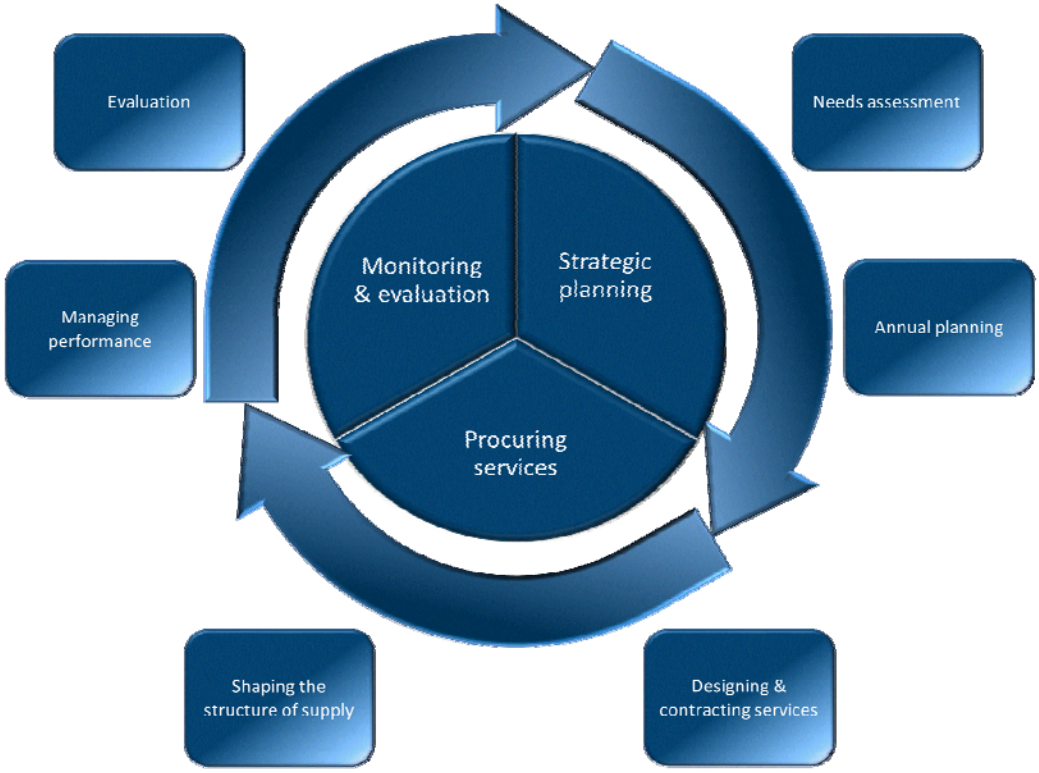
Useful information in this resource:

_____

Australian Government
Department of Health

phn COMMISSIONING
An Australian Government Initiative

Needs assessment guide



December 2015

Acknowledgement

The Australian Government Department of Health acknowledges and appreciates the input of Primary Health Networks in the development of this Guide.

Note

This guide does not override the requirements set out in the PHN Funding Agreement.

Contents

Figures and tables

1. THE PHN COMMISSIONING FRAMEWORK

Primary Health Networks (PHNs) were established in July 2015, with the objectives of:

- increasing the efficiency and effectiveness of medical services for patients, particularly those at risk of poor health outcomes; and
- improving coordination of care to ensure patients receive the right care in the right place at the right time.

The PHN Guidelines state that:

> PHNs will work towards achieving these objectives on the basis of an understanding of the health care needs of their communities through analysis and planning. They will do this through knowing what services are available and help to identify and address service gaps where needed, including in rural and remote areas, while getting value for money.[1]

Figure 1. The PHN Commissioning Framework

Figure 1 depicts the PHN Commissioning Framework. There are three phases in the cycle – strategic planning, procuring services, and monitoring and evaluation.[2]

This commissioning framework has been developed so that PHNs can ensure that their commissioning approach is consistent with the approach adopted for the programme as a whole and that the process results in consistent, comparable and measurable outputs and outcomes.

It is important to keep in mind that commissioning is a holistic approach to enable PHNs to work as strategic organisations at the system level. It is not merely a process. It is expected that PHNs may well be engaged in different parts of the cycle throughout the year (such as monitoring contracts). While PHNs are required to undertake a review and update of the needs assessment annually, in practice the needs assessment should be under continual review as new information, data and experience become available.

Commissioning

'Commissioning' is a continual and iterative cycle involving the development and implementation of services based on planning, procurement, monitoring, and evaluation. While a commissioning approach is used in a number of sectors other than health care it has been a key feature of the health system in the United Kingdom since the 1990s and is also a feature of health systems in New Zealand and the United States of America. Commissioning describes a broad set of linked activities, including needs assessment, priority setting, procurement through contracts, monitoring of service delivery, and review and evaluation.

A key characteristic of commissioning is that procuring or purchasing decisions occur within a broader conceptual framework. The difference between purchasing and commissioning in the health care context has been described as follows:

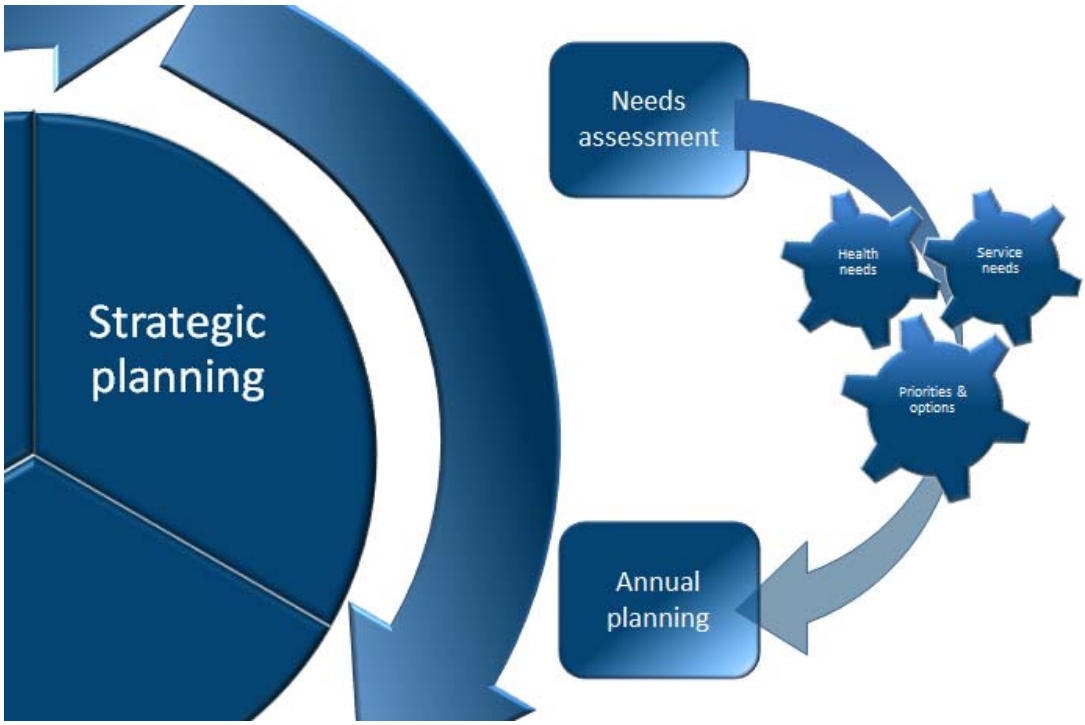
> Commissioning is a term used most in the UK context and tends to denote a proactive strategic role in planning, designing and implementing the range of services required, rather than a more passive purchasing role. A *commissioner* decides which services or healthcare interventions should be provided, who should provide them and how they should be paid for, and may work closely with the provider in implementing changes. A *purchaser* buys what is on offer or reimburses the provider on the basis of usage.[3]

As the health systems are different, PHN commissioning will of course differ from the experiences of other countries. However, the fundamental elements remain valid in the Australian context.

The Strategic Planning phase

Strategic planning is the first phase in the PHN Commissioning Framework. It consists of two stages – undertaking a needs assessment in order to identify and prioritise opportunities for activity, followed by the development of annual plans. The two components are closely linked, but distinct. Annual plans will be informed by factors other than the needs assessment such as cost, capacity and timing.

Figure 2. The Strategic Planning phase



Undertaking an analysis and assessment of the health and service needs of people in the PHN region enables the PHN to identify opportunities and set priorities for planning. These plans in turn shape the activities which enable it to achieve the PHN objectives.[4]

2. NEEDS ASSESSMENT

A Needs assessment is:

> a systematic method of identifying unmet health and healthcare needs of a population and making changes to meet these unmet needs. It involves an epidemiological and qualitative approach to determining priorities which incorporates clinical and cost effectiveness and patients' perspectives. This approach must balance clinical, ethical, and economic considerations of need—that is, what should be done, what can be done, and what can be afforded.[5]

Undertaking a needs assessment provides the PHN with the opportunity to engage with Local Hospital Networks (or equivalents) and other key planning and funding agencies in order to ensure alignment of effort and investment. This involves the identification and analysis of key data and other forms of information.

Opportunities for community empowerment will also begin in the needs assessment stage. Consulting communities is a key method for understanding factors which affect their health and quality of life, and is a means of recognising the needs of disadvantaged groups which may not be represented in routine statistical collections.[6] While a range of engagement approaches will need to be considered, Community Advisory Committees and Clinical Councils must play key roles in the development of the needs analysis.

As commissioning is ongoing and iterative, the needs assessments of future years will themselves be informed by the experience of previous years and by the learnings gained from the experience of monitoring and evaluating previous activities and investment.

Needs assessments should use existing data and evidence where possible and not duplicate the efforts of others, particularly Local Hospital Networks (or equivalents). Needs assessments should also:

- analyse relevant and current local and national health data including, but not limited to, data collected by Local Hospital Networks (or equivalent);
- review health service needs and available service provision in the region;
- identify health services priorities based on an in-depth understanding of the health care needs of the communities within the PHN region; and
- be informed by clinical and community consultation and market analysis.

The needs assessment will contribute to the development and implementation of an evidence-based Annual Plan to address national and PHN specific priorities relating to patient needs and service availability and gaps in the PHN region.

While it is important for the needs assessment to be systematic this does not mean it attempts to cover the entire scope of primary health care. The PHN needs assessment should focus on:

- the PHN objective of efficiency and effectiveness of medical services for patients, particularly those at risk of poor health outcomes;
- the PHN objective of opportunities to improve coordination; and
- the six key priorities for targeted work: mental health, Aboriginal and Torres Strait Islander health, population health, health workforce, eHealth and aged care.

PHNs should ensure that attention is given to the health needs of Indigenous Australians, recognising the commitment of all parties to Closing the Gap.

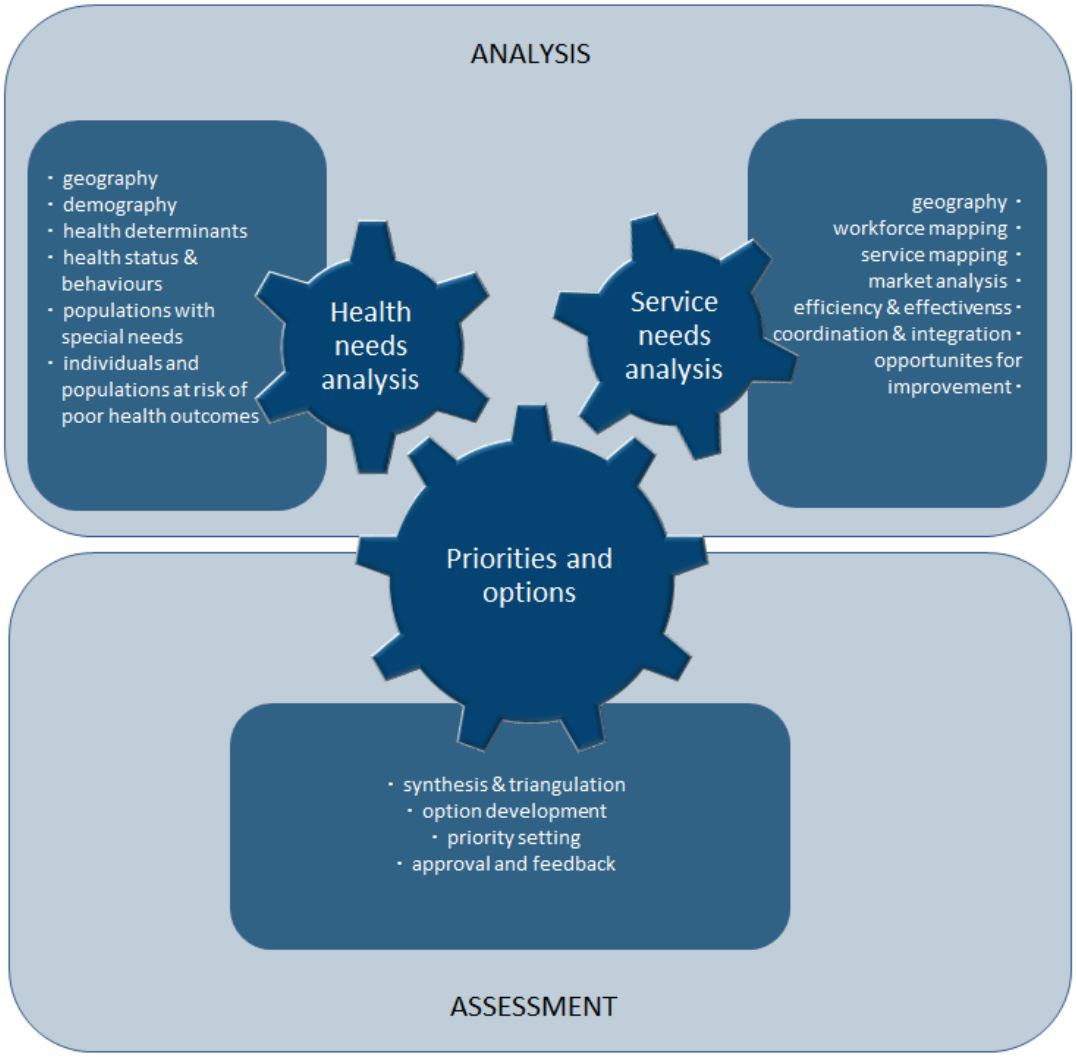
Structure

The PHN needs assessment consists of two parts: analysis and assessment (Figure 3).

The main paradigm for the first part is *analysis* – the examination and documentation of health needs and service needs within the region. The analysis of health needs and the service analysis are shown as separate. While health needs and service needs are interlinked, it is important to consider each independently in the first instance. This does not mean to suggest ignoring the very real relationships between health needs and the nature and capacity of health services, as of course the nature of service provision (or lack of) can impact on health status. Equally, the region's demography also shapes, over time, the distribution of services. It is expected that these two work streams will be undertaken concurrently and for each to inform the other. Putting these two perspectives together is largely undertaken in the *assessment* part.

The main paradigm for the second part is *assessment* – where the PHN exercises a level of judgement about relative priorities, considers a number of alternative options and makes sometimes quite difficult decisions. It is important to remember that the needs assessment is <u>not</u> a plan. The needs assessment concludes with the identification of opportunities, priorities and options. Proposals for how these are acted upon – which may be through direct investment in purchasing services or by other means – are part of the subsequent planning stage.

Figure 3. Structure of the needs assessment



In both *analysis* and *assessment* the focus moves progressively from the general to the specific. In the health and services needs analysis it is expected that the PHN will move from a general analysis of – for example – population or workforce distribution across the entire region and towards a focus on particular groups, locations or service types that appear to be emerging as potential priorities. A similar narrowing of focus will characterise the assessment, as the PHN moves to a position where it can identify opportunities, priorities and options.

Approach

Two broad kinds of information will be sourced throughout the needs assessment. The first is information and data (both qualitative and quantitative) from a wide range of sources. The second is from consultations with communities, health professionals and other stakeholders.

Data and other sources of information

The Department of Health (the department) will consult with PHNs and other stakeholders about data needs in general, and what might be provided through a national website. It is also expected that through undertaking a baseline needs assessment PHNs will identify and further clarify data needs, with a particular focus on the six key priority areas for targeted work.

It is recognised that each PHN will have differing and sometimes unique sources of information available to them, which may be in a variety of forms and in some cases subject to restrictions on its use. As a result, there is no requirement for PHNs to develop a standardised population health profile or description of the nature of the health system or health service provision in the PHN.

PHNs can use information from the previous Medicare Locals, including the Comprehensive Needs Analyses, data and other information available through corporate knowledge and stakeholder networks. Potential data sources should be evaluated against standard measures of quality, such as the ABS Data Quality Framework.[7]

A number of data sources may be useful for both health needs analysis and service needs analysis. For the baseline needs assessment it is expected that the PHN will make use of a wide range of sources, including but not limited to the following:

- Australian Bureau of Statistics (ABS) Census and Census-derived data on demographics, including the Socio Economic Indices for Areas (SEIFA) and profiles of health including the National Health Survey, the National Nutrition and Physical Activity Survey and the National Health Measures Survey, the National Aboriginal and Torres Strait Islander Health Survey, the National Aboriginal and Torres Strait Islander Nutrition and Physical Activity Survey, and the National Aboriginal and Torres Strait Islander Health Measures Survey
- a range of Australian Institute of Health and Welfare (AIHW) and National Health Performance Authority (NHPA) datasets and publications, including the METeOR metadata registry
- Medicare Benefits Schedule (MBS), Pharmaceutical Benefits Scheme (PBS) and Practice Incentives Programme (PIP) data
- aged care data (both residential and community based) such as Commonwealth Home Support Programme, or Department of Veterans Affairs (DVA) data
- mental health data such as the Access to Allied Psychological Services (ATAPS) data collection

- information on Indigenous data from a range of sources
- the Australian Childhood Immunisation Register
- the National Notifiable Diseases Surveillance System
- resources from the Royal Australian College of General Practitioners (RACGP)
- data from practices through clinical audit tools, and the Bettering the Evaluation and Care of Health (BEACH) data
- health workforce data
- State and Territory Health Department data
- data from Local Hospital Networks or equivalents (including individual acute and community care services)
- Local Government data
- information on the PHCRIS website
- information on the PHIDU website, such as the Social Health Atlas of Australia
- National Health Services Directory (NHSD) and Healthdirect
- there are also some mapping tools that may be useful to PHNs. [8]

Other sources of information that may inform needs assessment are being considered in the development of the PHN Performance Framework, during which the department will consult with PHNs and other stakeholders. The Performance Framework will encompass three tiers of performance. The *national tier* will centre on a small set of headline indicators to assess performance against health outcomes. The *local tier* will assess performance using local performance indicators selected by PHNs through formal planning processes. The *organisational tier* will focus on compliance with contractual arrangements in key areas of establishment, operations and activity.[9]

There are a number of other sources of information that PHNs can use in structuring their thinking about the needs assessment. Many of these are frameworks that have been developed to classify and organise data and indicators about health status, service delivery and opportunities for quality improvement.[10] Information on overall health performance frameworks and reporting is provided in the publication by the Australian Institute of Health Innovation, *Performance indicators used internationally to report publicly on healthcare organisations and local health systems*.[11]

Consultation

Different approaches may be necessary for PHNs to tailor their consultation to specific regional characteristics and community needs. Whatever the approach, PHNs should ensure:

- a systematic approach is in place for identifying target groups and stakeholders, particularly those in the six key priority areas for targeted work, and determining consultation methodologies;

- consultation and engagement is undertaken in a structured and informed manner;
- targeted processes are used as necessary to gather the views of vulnerable or hard to reach population groups and to engage with specific groups, including Indigenous Australians;
- robust processes exist to capture, analyse and synthesise information gained from consultations; and
- mechanisms are in place to provide feedback on the consultations to both participants and communities more generally.

Consultation occurs throughout the needs assessment. While it is expected that the bulk of consultation will occur as PHNs undertake the health and service needs analyses, it is also an important element in the final stage of assessing and prioritising need. [12]

Consultations with communities

Consultation with the community can provide qualitative insights into the health needs of the population that cannot be elicited from data alone. Community consultation is essential in order to obtain information about the perceived needs of local communities and individuals, insight into the experiences of patients, consumers and carers, and their perspectives on how primary health care should be improved and where it is already working well.

Consultations can help identify barriers to positive health and wellbeing, ascertain how satisfied the local community is with existing health service provision and identify which health service provision could be improved to better meet their needs.

Consultation needs to gather both general views and the views of vulnerable or hard to reach population groups and those with specific health issues and support needs, which may be limited by health literacy, perceptions or knowledge of what options are available. PHNs should also recognise cultural diversity within their region and design appropriate consultative approaches. Community Advisory Committees are central to any community consultation.

Consultations with health professionals, providers, funders and other stakeholders

The perspectives of health professionals, providers, funders and other stakeholders can differ from community and consumer views, and are equally important in informing the needs assessment. Consultation with these stakeholders is necessary to obtain their perspectives and advice on health and service issues and needs.

Clinical Councils are central to the consultation process, including in the identification of key stakeholders and possible consultation mechanisms. Consulting with these stakeholders will provide an important perspective in areas such as:

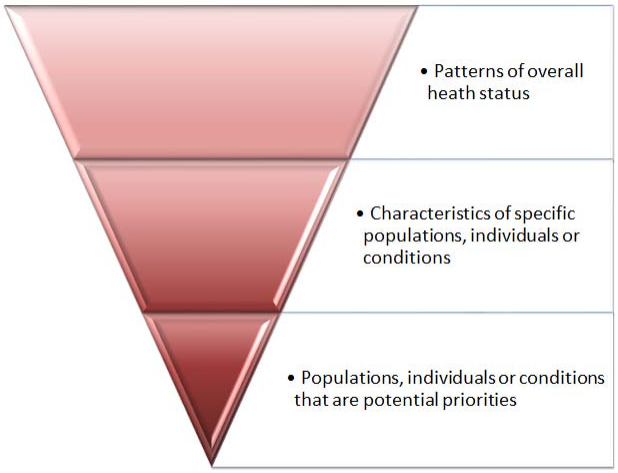
- the incidence and prevalence of health conditions and disease trends, including issues that influence health outcomes at the local level;

- workforce needs including capacity building and quality improvement, size, distribution and discipline mix, including issues and barriers at the local level;
- the efficiency and effectiveness of current service delivery models and opportunities for improvement;
- coordination and linkages between services and opportunities for improvement;
- the spread of services across the spectrum of prevention, diagnosis, early intervention, treatment, rehabilitation and continuing care; and
- vulnerable groups and barriers to access.

3. ANALYSIS

Health needs analysis

Figure 4. Health needs analysis



The intention of the health needs analysis is to ensure the PHN has an understanding of the health status and needs of individuals, populations and communities relevant to its role within both the health system and the broader environment.

The health needs analysis will need to make use of a range of demographic and epidemiological data, alongside structured consultations. It will also require some consideration of the wider social and economic determinants of health.

Figure 4 illustrates how the PHN moves through the various steps that make up this stage. The focus will progressively narrow towards an identification of people, populations or conditions that are likely to be priorities for the PHN.

The elements required to undertake a health needs analysis are discussed below. This is so that PHNs can ensure they cover all expected elements and in doing so progressively move towards identifying key priorities. In practice, many of these elements are inter-related or overlapping and the overall process will be iterative.

Geography

The Australian Statistical Geography Standard (ASGS) should form the basis of sub-regions the PHN develops for the purposes of demographic analysis, using the SA 3 level as the building block. Demographic data should be analysed at the SA 2 level where possible. The ASGS also includes the Indigenous Urban Centres and Localities and the Remoteness Area classifications. Depending on the data, there may be reasons for also using LHN sub-divisions or Local Government Areas. Use of classifications such as postcodes and electorates is generally discouraged.

Demography

Key demographic data will also be derived from the ABS Census or Estimates of Resident Population. Information on population trends can be significant, and the SEIFA is an important tool.

Health determinants

The nature and importance of various health determinants will vary from PHN to PHN. It is difficult to generalise about how each PHN will treat this area, but a guiding principle should be that the various areas under consideration are analysed in terms of their impact on health rather than issues in their own right. There are a number of frameworks and reference points, particularly the work of Wilkinson, Marmot, and the WHO.[13] Information on levels of health literacy in the PHN is also an important consideration.[14]

Health status and behaviours

At this point in the process the focus is on the PHN region as a whole. A range of morbidity and mortality data will need to be examined, alongside a process of consultation. This would include data from the Australian Health Survey and other ABS and AIHW surveys. Techniques for disaggregation or synthetic estimates at small area level need to be robust. As an example, the National Health Performance Framework sub-components are:

Health status

- deaths (mortality rates and life expectancy measures)
- health conditions (prevalence of disease, disorder, injury or trauma or other health-related states)
- human functions (alterations to body structure or function [impairment], activity limitations and restrictions on participation)
- wellbeing (measures of physical, mental and social wellbeing of individuals).

Health behaviours

- attitudes, beliefs, knowledge and behaviours such as patterns of eating, physical activity, smoking and alcohol consumption, and participation in cancer screening programmes .[15]

Where possible the analysis of health status should consider the make-up of the PHN region's total fatal and non-fatal burden in comparison with national and jurisdictional data, and whether or not this can be understood using measures such as disability-adjusted life years or other appropriate measures. While PHNs are expected to have a focus on chronic disease, the burden of communicable disease and injury needs to be considered.

Populations with special needs

This step may well cover much of the same ground as the work on health status, behaviours and determinants, but from a different starting point. In an influential article, the epidemiologist George Rose argued that:

> I find it increasingly helpful to distinguish between two kinds of aetiological question. The first seeks the cause of cases, and the second seeks the cause of incidence. 'Why do some individuals have hypertension?' is a quite different question from 'Why do some populations have much hypertension, whilst in others it is rare?' The questions required different kinds of study, and they have different answers.[16]

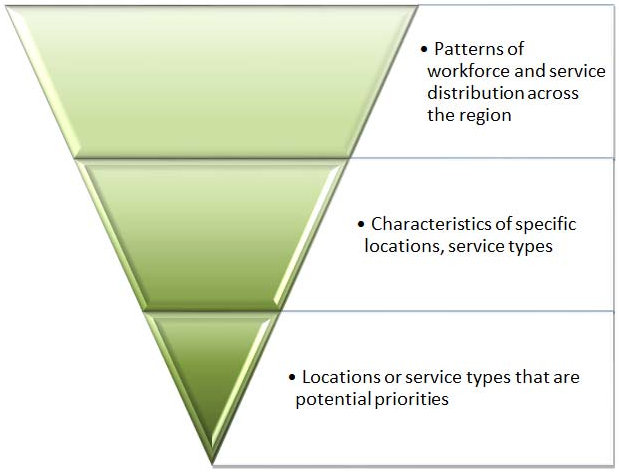
Many of the previous steps may have focused on the health conditions of individuals, albeit aggregated into groupings of one kind or another based on a number of premises. The needs assessment should involve an explicit consideration of populations with special needs that does not necessarily start from a health-related premise. This would involve identification of populations based on characteristics such as ethnicity, location, gender, age, socioeconomic status etc. and may identify issues or inequities specific to these (suicide in youth, injury in farming communities, etc.) which were less evident in the preceding analyses.

Individuals and groups at risk of poor health outcomes

This final step combines evidence from all of the above. Where possible, analysis may involve looking at relative measures of need such as disease rates in comparison to benchmarks or targets, national or jurisdictional averages, or to rates in other comparable regions.[17]

Service needs analysis

Figure 5. Service needs analysis



The intention of the service needs analysis is to ensure that the PHN has an understanding of the region's services and health infrastructure relevant to its role within both the health system and the broader environment.

An analysis of geography, workforce and services should provide each PHN with an overall picture of the nature of the health workforce and services in their region. In the latter parts of the analysis PHNs are asked to focus on service need issues from the perspective of the PHN objectives of efficiency, effectiveness and coordination.[18]

Figure 5 illustrates how the PHN moves through the various steps that make up this stage. The focus will progressively narrow towards an identification of specific locations, service types or relationships between services that are likely to be priorities for the PHN.

As with the health needs analysis, many of these elements are inter-related or overlapping and the overall process will be iterative.

Geography

There are a number of geographies that are essentially administrative or related to service delivery rather than demographic factors, including the boundaries of both PHNs and LHNs and jurisdictional borders. In analysing service need PHNs should consider how these are important in relation to the health system's capacity or performance at the regional level. This can include issues such as cross-border flows, distribution of services in adjacent PHN regions (particularly in urban areas), referrals both out of and into centralised services such as large teaching hospitals, specialist and allied health services, and the location of specialist imaging or diagnostic services. Variations in services provided by different local government authorities within a PHN region may also be relevant.

Workforce mapping

PHNs should look at information on the health workforce in the PHN region. If available, this could include:

- number and distribution by type, such as GPs, allied health, pharmacy, specialist services such as psychiatry, community health services etc.;
- characteristics such as full or part time, public versus private, qualified but not working in health care etc.; and
- relationships between professional groups.

Service mapping

Service mapping involves identifying and documenting the range of services available within the PHN region, and the kinds of relationships that exist between services. The needs assessment should consider aspects such as:

- ***location –*** including but not limited to physical location, hours of opening, with consideration of identifiable gaps. For rural areas this would also include outreach services, for urban areas it may involve some consideration of services outside the PHN borders but accessed by people from within the PHN;
- ***utilisation –*** including MBS and PBS data, a range of hospital data, such as use of Emergency Departments and Potentially Preventable Hospitalisations, other measures of occasion of service, and with a consideration of under-utilisation, duplication and waste;
- ***accessibility –*** including financial and cultural barriers and access to specialists and secondary referred services, access to services after hours;
- ***responsiveness –*** such as wait times;
- ***capability –*** such as skills and competence;
- ***acceptability –*** such as cultural sensitivity, patient experience of and satisfaction with the quality of care; and
- ***quality –*** such as practice accreditation and PIP enrolment.

Service mapping can also include consideration of the system's ability to deal with public health emergencies (such as an influenza pandemic) and issues around regional coordination that may impact on emergency preparedness.

Market analysis

As part of its focus on the supply side, the PHN should be alert to evidence on how the health market works in its area. This includes thinking about parts of the market not currently active in health care, but where there are potential opportunities for engagement. This is not necessarily limited to specific suppliers, but could involve services such as informatics or business models from other sectors.

Efficiency and effectiveness of health services

In this step, the PHN should consider service provision in terms of different concepts of efficiency:

- technical efficiency – the cost at which services are produced;
- allocative efficiency – how services reflect consumer preference, from a given set of resources; and
- dynamic efficiency and sustainability – the capacity of the system to sustain workforce and infrastructure, to innovate and respond to emerging needs. [19]

In terms of effectiveness, PHNs should consider accepted dimensions of effectiveness. The AIHW defines effectiveness as how well the outputs of a service achieve the stated objectives of that service. Indicators to measure this will cover characteristics such as access, quality and appropriateness, including issues such as cultural competency of services for Indigenous and CALD communities.[20] [21]

An analysis in terms of equity is also important. A focus on programmes and the way they are funded, rather than on what the patient needs, allocates resources inequitably as well as inefficiently.[22] Horizontal equity is exhibited when services are equally accessible to everyone in the community with a similar level of need, and vertical equity is exhibited when it accounts for the special needs of certain groups in the community and adjusts aspects of service delivery to suit these needs.[23] [24] [25]

Coordination between and integration of, services

PHNs should analyse the level of coordination and integration of health care services in the region, where opportunities may exist for improvement and the presence or absence of services that seek to directly address coordination. Consideration could include evidence in areas such as:

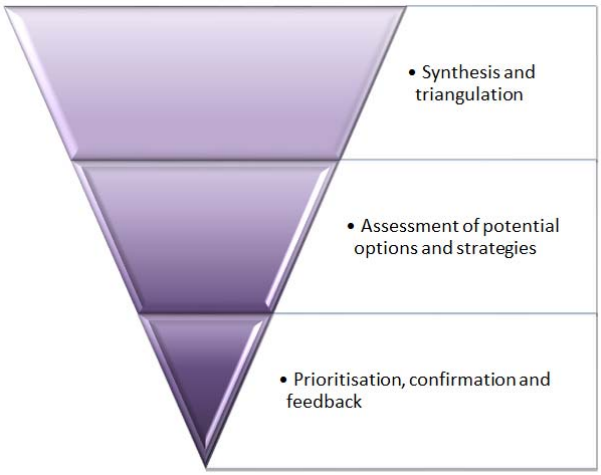
- shared health records and other e-health initiatives;
- examples of integrated service delivery;
- models such as the Patient Centred Medical Home;
- transitions between acute care and primary care;
- coordination between general practice and allied health;
- linkages between health and social services (aged care, disability services, youth, child and family services, housing); and
- referral patterns and health pathways.

Strengths and weaknesses

This final step combines evidence from all of the above. It is acknowledged that in some areas the baseline 2015-16 needs analysis will be preliminary. In the first iteration PHNs may not have developed a comprehensive understanding of the overall health market, or fully developed ways to assess service need from a perspective of efficiency, effectiveness and coordination.

4. ASSESSMENT

Figure 6. Assessment, priorities and options



Assessment has the following purposes:

- to *synthesise* and *triangulate* evidence from consultations and the analysis of the health and service need components; and
- on the basis of this, identify opportunities for further consideration; and
- determine *priorities* and assess *options* for further development in the PHN Annual Plan.

When considering options, it is expected that at this point the PHN would be seeking information from a variety of sources (including literature and systematic reviews) on the nature of interventions and approaches that have been implemented elsewhere, evidence of their success and appropriateness for use in the PHN. Some of this evidence may have emerged during the first two stages.

Figure 6 illustrates how the PHN moves through the various steps that make up this stage. The focus will progressively narrow towards an identification of people, populations or conditions that are likely to be priorities for the PHN and identify options to address these.

Synthesis and triangulation

Issues and needs arising from the data or identified through the community, professional and stakeholder consultations will have been summarised into consistent themes progressively. This information and evidence now needs to be compared and cross checked.

Triangulation is the use of more than one method of enquiry to assess and verify findings. It is used to bring the results of the qualitative and quantitative analyses together and to confirm major themes and key issues identified through the needs assessment process. Here triangulation can be used to cross check, confirm and/or verify the issues identified through community and stakeholder consultations with the findings of the analyses of data or service utilisation patterns.

The matrix below illustrates a simple method to compare how issues raised in community or stakeholder consultations can be cross checked with health needs data or service usage information. A simple star scoring method can be used to assess and compare a list of health and service needs generated from the consultations and data analyses.[26]

Figure 7. Triangulation matrix

Issue	Community/ consumer feedback	Service provider feedback	Health needs analysis	Service needs analysis	Triangulation result
Health Issue					
Service issue					

Priority setting and options

Prioritisation, or priority setting, applies to the full range of opportunities for PHN activities, not just new or marginal ones. This involves an assessment of whether current service arrangements are best suited to meeting identified need. Whatever processes are adopted for prioritisation, PHNs need to ensure that:

- they are evidence-based;
- are balanced and take account of the views of different groups and parties; and
- decision-making processes are transparent, fair and reasonable.[27] [28]

It is expected that as the needs assessment progresses the PHN will be identifying the kinds of interventions, programmes or policies that have been developed in the past to address a number of the identified health and service needs. These should now be assessed as possible options on the basis of their likely appropriateness (including funding available, effectiveness, efficiency and value for money).

The existence or otherwise of evidence based models will inevitably impact on priority setting. An issue may be identified as a priority, but with no obvious way forward the PHN may need instead to invest in further development, which is consistent with its longer term role in shaping the market.

5. SUMMARISING THE FINDINGS

Developing a summary of the findings of the health needs analysis, service needs analysis and the priority setting process will help to inform PHN Annual Plans and facilitate reporting and information sharing. In addition, the department may use this information to inform programme and policy development.

The following tables illustrate how key information from the needs assessment can be provided in a simple and common format. This does not prevent PHNs using the needs assessment process to develop more detailed or different outputs (such as population health profiles) as circumstances require.

PHNs will report to the Department of Health on their needs assessment activity.

Table 1. Outcomes of the health needs analysis

This table illustrates how findings of the health needs analysis can be summarised.

Identified Need	Key Issue	Description of Evidence
e.g. *Health status*	*Poor self- assessed health status in (specific locations)*	*NHPA analysis of ABS Patient Experience Survey 2013–14.Variation within PHN obtained through consultations. Utilisation of hospitals and general practices.*
e.g. *Chronic disease*	*Lifestyle and risk factors impact on the development of (identified) chronic diseases*	*Variations between population groups and locations. Evidence from community consultations on diet, smoking, drug and alcohol use, physical activity and health determinants. Hospitalisations for chronic diseases.*

Table 2. Outcomes of the service needs analysis

This table illustrates how findings of the service needs analysis can be summarised.

Identified Need	Key Issue	Description of Evidence
e.g. *Service coordination*	*Lack of easily understood and accessible referral pathways across settings and providers*	*Proportion of GPs, community based medical specialists and allied health providers with secure messaging systems. Discharge information, analysis of directory and eligibility criteria for services, awareness of services from consultations with consumers and health professionals.*
e.g. *Mental health*	*Lack of child and adolescent psychiatry services in specific locations.*	*Identified through consultation with mental health workers and school counsellors.* *Data on available psychiatric sessions per capita, waiting list for assessment and caseload information.*

Table 3. Opportunities, priorities and options

This table illustrates how PHNs can summarise the priorities arising from the Needs Assessment and options for how they will be addressed. This could include options and priorities that:

- may be considered in the development of the PHN Annual Plan, and supported by PHN flexible funding;
- may be undertaken using programme-specific funding; and
- may be led or undertaken by another agency.

Priority	Possible options	Expected outcome	Possible performance measurement	Potential lead

6. ANNUAL PLANNING

The Needs Assessment component has a definite endpoint with the identification of opportunities, priorities and options. Decisions about which activities will be undertaken to take these forward are the concern of the annual planning – the second component of the strategic planning phase.

In their Annual Plans, PHNs may pursue a number of priorities through flexible funding. In addition, some priorities may more appropriately be addressed through the use of separate programme-specific funding. While an annual plan has a particular focus on the upcoming financial year it will also include medium and longer term perspective.

A number of areas may be identified where further investigation is required. There may also be identified priorities where it will take longer to develop responses, or where it is more appropriate that another agency such as Local Hospital Networks or equivalents would take the lead role.

ENDNOTES

[1] Australian Government Department of Health. *Primary Health Networks Grant Programme Guidelines* 2014. p.7.

[2] The commissioning cycle is most commonly presented in a diagram. This diagram is based largely on that developed by the NHS Information Centre and used to support World Class Commissioning between 2000 and 2010. See the NHS Information Centre archive and Commissioning Handbook for Librarians. Another model that is used extensively, with variations, was developed by the Institute for Public Care. First developed in 2003 and since adapted by a number of different agencies, the IPC cycle shows the relationship between strategic commissioning (the outer circle) and procurement, contracting and purchasing (the inner circle). This model follows the 4 step Plan-Do-Study-Act cycle first developed by Deming and used as the basis for many quality control and continuous improvement programmes. See: Institute of Public Care. *Commissioning for Health and Social Care*. Oxford Brookes University 2014, pp.11-13. Bovaird T et al. *Commissioning across government: review of evidence*. Third Sector Research Centre Research Report 86: for the UK National Audit Office. August 2012, pp.48-49.

[3] Smith J, Curry N, Mays N, Dixon J. *Where next for commissioning in the England NHS?* The Nuffield Trust and the King's Fund 2010, p.12. Also see Øvretveit J. *Purchasing for health: a multidisciplinary introduction to the theory and practice of health purchasing*. Open University Press 1995.

[4] Australian Department of Health. *Primary Health Networks Grant Programme Guidelines* 2014.p.10.

> Commissioning is characterised by a strategic approach to procurement that is informed by the baseline needs assessment and associated market analysis undertaken in 2015-16. Commissioning will enable a more holistic approach in which PHNs can plan and contract medical and health care services that are appropriate and relevant to the needs of their communities. Commissioning is further characterised by ongoing assessment to monitor the quality of services and ensure that relevant contractual standards are fulfilled. It is expected that PHN commissioning capabilities will continue to develop over time.

[5] Wright J, Williams R, Wilkinson JR. *Development and importance of health needs assessment*. BMJ 1998: 316: 1310-1313. Also see Katterl R et al. *Regionally-based needs assessment in Australian primary health care.* PHCRIS 2011.

[6] Smith B, Kwok Cho Tang, Nutbeam D. *WHO Health Promotion Glossary: new terms.* Oxford University Press 2006.

[7] ABS Data Quality Framework Cat No. 1520.0. 2009 and National Statistics Service: Data Quality Online. The ABS DQF is comprised of seven dimensions of quality, reflecting a broad and inclusive approach to quality definition and assessment. The seven dimensions of quality are *Institutional Environment, Relevance, Timeliness, Accuracy, Coherence, Interpretability and Accessibility***.** All seven dimensions should be included for the purpose of quality assessment and reporting. However, the seven dimensions are not necessarily equally weighted, as the importance of each dimension may vary depending on the data source and context.

[8] See the following.

ABS	ABS website - Census homepage
	ABS website - census - SEIFA
	ABS website - Australian Health Survey
	ABS website - Profiles of Health, Australia 2011-13
	ABS website - Australian Aboriginal and Torres Strait Islander Health Survey: Physical activity, 2012-13
AIHW	AIHW website - Homepage
	AIHW website - Metadata Online Registry homepage
NHPA	NHPA website - Homepage
Medicare	Department of Health website - MBS online
	Department of Human Services website - Medicare item statistics/reports
ACIR	Department of Human Service website - Australian Childhood Immunisation Register for health professionals
Notifiable Diseases	Department of Health website - National Notifiable Diseases Surveillance System
	Department of Health website - National Notifiable Diseases Surveillance System Annual report
RACGP	RACGP website - Homepage
Indigenous health	Department of Health website - Aboriginal and Torres Strait Islander Health Performance Framework
	Australian Indigenous Health InfoNet website - Homepage
DVA	Department of Veterans Affairs website - Data and Statistics - Statistics about the veteran population
Mental health	Access to Allied Psychological Services website - Minimum dataset
	ABS website - Mental Health Statistics
	AIHW website - Mental Health
Health workforce	Health Workforce Australia website - History
	Health Workforce Australia website - National Statistical Resource
	AIHW website - Workforce
	AIHW website - Workforce publications
BEACH	University of Sydney website - Medicine Research Centre - Bettering the Evaluation and Care of Health
PCHRIS	Primary Health Care Research and Information Service website - homepage
PHIDU	University of Adelaide website - Public Health Information Development Unit
	University of Adelaide website - Public Health Information Development Unit- Data
APHCRI	Australian National University website - National Centre for Geographic and Resource Analysis in Primary Health Care
NHSD	National Health Services Directory website - Homepage
Healthfirst	Health First Network website - Homepage

[9] PHN Performance Framework

[10] See the following:

- National Health Information Standards and Statistics Committee (NHISSC). *The National Health Performance Framework* 2nd Edition 2009. the Australian Health Performance Framework consists of three domains; health status, determinants of health and health system performance. Click

here to go to AIHW website - Metadata Online Registry - National Health Performance Framework. This performance framework is used (in a slightly modified form) to monitor progress in Indigenous Australian health outcomes, health system performance and broader determinants of health. Australian Health Ministers' Advisory Council. *Aboriginal and Torres Strait Islander Health Performance Framework 2014 Report*. AHMAC 2015. This is the 5th in this series which have been released every two years since 2006. Click here to go to the Department of Health website - Aboriginal and Torres Strait Islander Health Performance Framework

- Productivity Commission. *Report on Government Services* 2014. Chapter 1: Approaches to performance measurement. This framework groups indicators under three broad headings of equity, effectiveness and efficiency, with access as a subset of both equity and effectiveness. Note that this framework is specifically designed to report on government services. Click here to go to the Australian Government Productivity Commission website - Report on Government Services
- National Health Performance Authority. *Performance and Accountability Framework* 2011. The Medicare Local component was based on the ROGS and proposed 31 initial indicators for Medicare Locals. Click here to go to the National Health Performance Network website - Performance Indicator Reporting
- CIHI. *A performance measurement framework for the Canadian health system.* Canadian Institute for Health Information 2012; IBM. *Evaluation of the Health Information Roadmap Initiative: Roadmap II and Roadmap II Plus.* 2007. The Canadian Health Roadmap has four dimensions: health status, non-medical determinants of health, health system performance, and community and health system characteristics.
- *The Triple Aim* considers health care in terms of improving the health of populations, improving the individual experience of care, and reducing the per capita costs of care for populations. This model is a key element in performance measurement in many US health care organisations – particularly since the Affordable Care Act – and is being adopted in a number of countries including New Zealand, the UK and Canada. Click here to go to the Institute for Healthcare Improvement website - Homepage

[11] Hibbert, P., Hannaford, N., Long, J., Plumb, J. and Braithwaite, J. *Final Report: Performance indicators used internationally to report publicly on healthcare organisations and local health systems.* Australian Institute of Health Innovation, University of New South Wales 2013.

[12] The International Association for Public Participation's core values, code of ethics, and public participation spectrum are a useful foundation for informing stakeholder engagement. See: Click here to go to the International Association for Public Participation Australasia website - Homepage

[13] See the following:

- Wilkinson RG, Marmot M. *The Solid Facts: the social determinants of health* 2nd edition. International Centre for Health and Society, WHO 2003. Wilkinson RG. *Socioeconomic determinants of health: Health inequalities: relative or absolute material standards?* BMJ 1997: 314:591. Wilkinson RG, Pickett KE. *Income inequality and population health: a review and explanation of the evidence.* Social Science and Medicine 2006: 62: 1768-1784.
- Marmot M. *Social determinants of health.* Oxford University Press 2005. Marmot M et al. *Closing the gap in a generation: health equity through action on the social determinants of health.* The Lancet 2008: 372: 1662-1669. See the Institute of Health Equity for more information and links.
- WHO. *Closing the gap in a generation: Health equity through action on the social determinants of health*. WHO. Final Report of the Commission on Social Determinants of Health 2008. 2011 World Conference on Social Determinants of Health.

[14] ACSQHC. *Health literacy: taking action to improve safety and quality.* Australian Commission on Safety and Quality in Healthcare 2015. p.2.

[15] See endnote 11 above. The National Health Performance Framework locates 'health behaviours' within the Health Determinants rather than the Health Status domain. The Aboriginal and Torres Strait Islander Health Performance Framework uses the sub components of 'health conditions', 'human function', 'life expectancy and wellbeing' and 'deaths'.

[16] Rose G. *Sick individuals and sick populations*. International Journal of Epidemiology 2001: 30: 427-432. First published International Epidemiological Association 1985. p.428. See also Merchant A et al. *Insights from cross-population studies: Rose revisited*. International Journal of Epidemiology 2005: 34: 344-246 [commentary].

[17] The National Health Performance Healthy Communities reports are an example of comparisons between 'peer groupings' of Medicare Locals. A range of methodologies are found in Begg S, Vos T, Barker B, Stevenson C, Stanley L, Lopez AD. *The burden of disease and injury in Australia 2003*. PHE 82. Canberra: AIHW 2007. The AIHW is currently undertaking the third Australian burden of Disease Study using 2011 data with results expected in the first half of 2016. Click here to go to the AIHW website - Burden of Disease

[18] The PHN may wish to consider the effectiveness, efficiency and coordination of services within the scope of a broader description of quality. There are many examples of these and include the work of Donabedian, the Australian Commission on Safety and Quality in Health Care, the WHO or the US Institute of Medicine. See the following:

- Donabedian A. *The seven pillars of quality.* Archives of Pathology and Laboratory Medicine 1990: 114: 11:1115-1118, and Donabedian A. *Evaluating the quality of medical care.* Milbank Quarterly 1966: 44: 3 Pt 2. Reprinted 2005: 83: 4: 691-729.
- Australian Commission on Safety and Quality in Health Care. *Australian Safety and Quality Framework for Health Care.* 2012
- WHO. *Quality of care: A process for making strategic choices in health systems.* World Health Organisation 2006. pp.9-10.
- Institute of Medicine. Crossing the Quality chasm: a new health system for the 21st century US IOM 2001. See also ARHQ. National Quality Measures Clearinghouse - Domain Framework. The US Department of Health and Human Services Agency for Healthcare Research and Quality (AHRQ) has a 'domain framework' which takes the definition of quality to another level.

[19] AIHW website - Metadata Online Registry - Report on Government Services - Effectiveness and AIHW website - Metadata Online Registry - Report on Government Services - Efficiency and Sustainability

[20] AIHW website - Metadata Online Registry - Report on Government Services - Effectiveness [summarised]

- *Access indicators* measure how easily the community can obtain a service. Access has two main dimensions, undue delay (timeliness) and undue cost (affordability). Timeliness indicators can include waiting times (for example, in public hospitals and for aged care services). Affordability indicators relate to the proportion of income spent on particular services (for example, out-of-pocket expenses in children's services).
- *Appropriateness indicators* measure how well services meet client needs. Appropriateness indicators also seek to identify the extent of any underservicing or overservicing. Data on differences in service levels can indicate where further work could identify possible underservicing or overservicing.

- *Quality indicators* reflect the extent to which a service is suited to its purpose and conforms to specifications. There is usually more than one way in which to deliver a service, and each alternative has different implications for both cost and quality. Information about quality is needed to ensure all relevant aspects of performance are considered.

[21] Australian Government. *Aboriginal and Torres Strait Islander Health Performance Framework 2014 Report*. p.142.

> Improving the cultural competency of health care services can increase Aboriginal and Torres Strait Islander peoples' access to health care, increase the effectiveness of care that is received, and improve the disparities in health outcome. Cultural competency requires that organisations have a defined set of values and principles, and demonstrate behaviours, attitudes, policies and structures that enable them to work effectively cross-culturally.

[22] Australian Government. *Reform of the Federation White Paper: roles and responsibilities in health.* Issues paper 3. Dec 2014. Click here to go to the Australian Government Reform of the Federation White Paper website - Roles and Responsibilities in Health Issues Paper Three - Equity

[23] AIHW website - Metadata Online Registry - Report on Government Services - Effectiveness

[24] Duckett S, Willcox S. *The Australian Health Care System* 4th edition. Oxford University Press 2011. pp.304-306. The areas used as overall headings for evaluating health care systems are equity, quality, efficiency and acceptability. Duckett and Willcox see effectiveness as a component of allocative efficiency, along with technical efficiency and priority setting. In this dynamic efficiency is conceptualised a bit differently and concerns the extent to which the health care system as a whole, and its constituent elements, adapt to change and innovation.

[25] World Health Organisation. Health Systems Financing: the path to universal coverage 2010 p.61.

> Efficiency… is a measure of the quality and/or quantity of output (i.e. health outcomes or services) for a given level of input (i.e. cost). So efficiency gains could help to contain costs – an important objective in many countries – by reducing the costs of service delivery. However, no one wants to contain costs by reducing health outcomes, so seeking efficiency gains should also be seen as a means of extending coverage for the same cost.

[26] This is a shorter version adapted from the Medicare Local Comprehensive Needs Assessment Tools and Resources, Appendix K. The star scoring system proposed was:

* rarely raised as an issue/not evident in data
** raised as an issue/somewhat evident in data
*** raised frequently as an important issue/concern evident in data
**** raised frequently as a high priority issue or concern needing action/significant concern evident from data

[27] Robinson S, Dickenson H, Williams I, Freeman T, Rumbold B, Spence K. *Setting priorities in health: a study of English primary care trusts.* The Nuffield Trust 2011, p.13.

[28] NHS Institute for Innovation and Improvement 2012. Prioritise Commissioning Opportunities
Also see the Qualitas Consortium and the Commissioning Handbook for Librarians

Research on Social Work Practice
21(5) 539-548

DOI: 10.1177/1049731511404436
http://rsw.sagepub.com

Process Evaluation of a Positive Youth Development Program: Project P.A.T.H.S.

Ben M. F. Law[1] and Daniel T. L. Shek[2,3,4]

Abstract

There are only a few process evaluation studies on positive youth development programs, particularly in the Chinese context. Objectives: This study aims to examine the quality of implementation of a positive youth development program (Project Positive Adolescent Training through Holistic Social Programs [P.A.T.H.S.]) and investigate the relationships among program adherence, process factors, implementation quality, and success. Method: Process evaluation of 20 Secondary 3 classroom-based programs was conducted in 14 schools. Results: Overall program adherence, individual evaluation items, quality, and success had high ratings. Principal components analysis showed that two components, namely, implementation process and implementation context were extracted from 11 evaluation items. The correlational analysis indicated that program adherence, implementation process, and context were highly correlated with quality and success. Multiple regression analyses show that teaching process and program adherence predicted quality, whereas teaching process, teaching context, and program adherence predicted success. Conclusions: The implementation quality of the Tier 1 Program of Project P.A.T.H.S. was generally high.

Keywords

Project P.A.T.H.S, process evaluation, positive youth development program

Introduction

Social work programs are specific sets of strategies and actions that can be implemented to enhance social functioning and problem-solving capabilities among individuals, families, and groups. Program evaluation is a systematic assessment of the process and outcomes of the programs with the aim of contributing to the improvement of the programs, such as in deciding whether to adopt the program further, enhancement of existing intervention protocols, and compliance with a set of explicit or implicit standards (Zakrzewski, Steven, & Ricketts, 2009). This article documents the process evaluation of a large-scale positive youth development program in Hong Kong called Positive Adolescent Training through Holistic Social Programs (P.A.T.H.S.).

Process Evaluation in Prevention Science and Social Work Practice

Outcome evaluation focuses mainly on the results of the programs, whereas process evaluation is concerned with how the program is actually delivered (Dane & Schneider, 1998; Domitrovich & Greenberg, 2000). Process evaluation is widely adopted in prevention science, such as nursing care (Huryk, 2010; Painter et al., 2010), chronic illness prevention programs (Braun et al., 2010; Karwalajtys et al., 2009; Mair, Hiscock, & Beaton, 2008; Shevil & Finlayson, 2009), smoking cessation programs (Gnich, Sheehy, Amos, Bitel, & Platt, 2008; Kwong et al., 2009; Quintiliani, Yang, & Sorensen, 2010), dietary programs (Allicock et al., 2010; Bowes, Marquis, Young, Holowaty, & Issac, 2009; Hart et al., 2009; Muckelbuer, Libuda, Clausen, & Kersting, 2009; Salmela, Poskiparta, Kasila, Vahasarja, & Vanhala, 2009), and AIDS rehabilitation programs (Bertens, Eiling, van den Borne, & Schaalma, 2009; Fraze et al., 2009; Hargreaves et al., 2009; Konle-Parker, Erien, & Dubbert, 2010; Mukoma et al., 2009). In social work practice, process evaluation has been used in family programs (Cohen, Glynn, Hamilton, & Young, 2010; Kumpfer, Pinyuchon, de Melo, & Whiteside, 2008) but is not commonly used in youth programs (Beets et al., 2008; Frazen, Morrel-Samuels,

[1] Department of Social Work and Social Administration, The University of Hong Kong, Hong Kong
[2] Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong
[3] Public Policy Research Institute, The Hong Kong Polytechnic University, Hong Kong
[4] College of Medicine, University of Kentucky, KY, USA,

Corresponding Author:
Ben M. F. Law, Department of Social Work, The Chinese University of Hong Kong, Hong Kong
Email: BenLaw@hku.hk

Reischl, & Zimmerman, 2009; Johnson, Lai, Rice, Rose, & Webber, 2010).

Process evaluation consists of five components, namely, program adherence, implementation process, intended dosage, macro-level implication, and process-outcome linkage (Scheirer, 1994).

Program adherence deals with whether the program is being delivered as intended according to the original program design. It is an important factor affecting the quality of program implementation (Domitrovich & Greenberg, 2000; Fagan, Hanson, Hawkins, & Arthur, 2008). True program fidelity is not easily achieved because program implementers often change or adapt the program content during actual implementation, whether intentionally or otherwise. Studies have shown that a number of preventive programs do not follow the prescribed program content entirely, and adaptation is made to specific target groups (Elliot & Mihalic, 2004; Nation et al., 2003). A study has found tension between the desire of the program implementer to adhere to the manualized plan and to make adaptations in accordance with the needs of clients (Wegner, Flisher, Caldwell, Vergnani, & Smith, 2008). Although it is not an easily resolved issue, program fidelity is generally encouraged, especially when programs are designed with vigorous trial runs and repeated success rates (Griffin et al., 2010; Johnson et al., 2010; Wilson et al., 2009).

Process factors are those that can be observed during the implementation process and are contingent to implementation quality or success. There is a variety of process factors according to the program characteristics and the needs of program developers. Some programs even design their own process measurements (Yamada, Stevens, Sidani, Watt-Watson, & de Silva, 2009). There are two main groups of process indicators in prevention science and social work programs. First is the implementation process. It is the direct observation of the interaction between the program implementer and the program receivers, such as the program receivers' engagement and the program implementer's use of feedback. Second is the implementation context. It involves context factors critical to implementation, including goal attainment and background knowledge, such as the program implementer's familiarity with the program receivers and the program implementer's program preparation.

Program dosage refers to the effort by program implementers to follow the required time prescribed for a program, as inadequate time affects the quality of program implementation (Bowes et al., 2009; Johnson et al., 2010). Dosage also refers to the group size of program receivers. A discrepancy between the intended and actual program receiver to program implementer ratio affects the program delivery process (Frazen et al., 2009).

Process evaluation can provide important findings with macro-level program implications, such as the importance of engagement of different community stakeholders (Carswell, Hanlon, O'Grady, Watts, & Pothong, 2009; Zani & Cicognani, 2010), client needs (Kwong et al., 2009), assessment of the environment (Eisenberg, 2009; Stewart, 2008), and challenges of the programs for a particular context (Louis et al., 2008).

Process evaluation and outcome evaluation are strongly linked. Process evaluation sheds light on which types of interventions strategies or process are related to the program success (Kwong et al., 2009; Painter et al., 2010). These factors can be amplified during program reimplementation.

The components of process evaluation point toward its importance. First, outcome evaluation provides inadequate hints on the quality of program implementation. Process evaluation demystifies the "black box" of intervention and aids in the understanding of the elements of program success or failure (Harachi, Abbot, Catalnao, Haggerty, & Fleming, 1999). Process evaluation facilitates program developers to understand fully the strengths and weaknesses of the developed programs. Program implementers can follow the suggestions from the process evaluation for further program delivery. This is one essence of evidence-based practice. It is also the foundation of bridging the gap between research and practice (Saul et al., 2008; Wandersman et al., 2008). Second, process evaluation can inform program developers about whether the programs are delivered according to some standardized manuals. The existence of other activities different from those intended by the program developers will not truly reflect the effectiveness of the prescribed programs. Third, different human organizations and communities arrange the programs in various settings, levels of involvement by the stakeholders, perceptions of the program among program implementers and program receivers, as well as the levels of support. Process evaluation can document the variety of implementations in real human service settings for the same manualized plans. Finally, process evaluation provides insights for program developers and implementers into the linkage between process and outcome. These insights allow both program developers and implementers to delineate the success and improvement areas during the process and connect them with the program outcomes.

Project P.A.T.H.S.

Many primary prevention programs and positive youth development programs have been developed in the West to address the growing adolescent development problems, such as substance abuse, mental health problems, and school violence (Shek, 2006a; Shek & Merrick, 2009). However, in Hong Kong, there are very few systematic and multiyear positive youth development programs. To promote holistic development among adolescents in Hong Kong, The Hong Kong Jockey Club Charities Trust approved the release of HK$750 million (HK$400 million for the first phase and HK$350 million for the second phase) to launch a project entitled "P.A.T.H.S. to Adulthood: A Jockey Club Youth Enhancement Scheme." The acronym "P.A.T.H.S." denotes Positive Adolescent Training through Holistic Social Programs. The Trust invited academics from five universities in Hong Kong to form a research team to develop a multiyear universal positive youth program (Shek & Merrick, 2009).

The project commenced in 2004 and is targeted to end by 2012. There are two tiers of programs in this project. The Tier

1 Program is a universal positive youth development program where students from the Secondary 1 (Grade 7) to Secondary 3 (Grade 9) participate in a classroom-based program, normally with 20 hr of training in the school year in each grade. Around one fifths of adolescents with more psychosocial needs will join the Tier 2 Program. The Tier 2 Program consists of intensive training on volunteer service, adventure-based counseling camp, and other experiential learning activities.

The overall objective of the Tier 1 Program is to promote holistic development among junior secondary school students in Hong Kong. The programs are designed according to 15 constructs conducive to adolescent development (Shek, 2006b): promotion of bonding, cultivation of resilience, promotion of social competence, promotion of emotional competence, promotion of cognitive competence, promotion of behavioral competence, promotion of moral competence, cultivation of self-determination, promotion of spirituality, development of self-efficacy, development of a clear and positive identity, promotion of beliefs in the future, provision of recognition for positive behavior, provision of opportunities for prosocial involvement, and promotion of prosocial norms.

The Tier 1 Program has several characteristics. First, there are 40 units per grade (each lasting for 30 min), with a total of 120 units, for the entire Tier 1 program. The time fits well with the Hong Kong secondary school time slots. Second, each school can choose to implement all 40 units (full program) or 20 units (core program), according to school needs. Third, the program content was developed by the research team and underwent extensive integration of existing research findings, adolescent needs, cultural characteristics, and trial teaching runs. Fourth, relevant adolescent developmental issues, such as drug issues, sexuality, and financial management, are incorporated into the program content so that it fits the current real-life experiences of Hong Kong adolescents. Fifth, the program implementers are either social workers or teachers who had to undergo intensive 20-hr training before program delivery.

There are two implementation phases: the experimental implementation phase (EIP) and the full implementation phase (FIP). The EIP aims at accumulating experience from trial teaching and administrative arrangement. Program materials are revised and refined during this phase. The FIP aims at executing the programs in full force. There are several lines of evidence that support the effectiveness of the Tier 1 Program, including the evaluation findings based on randomized group trials (e.g., Shek & Ma, 2011; Shek & Yu, 2011), subjective outcome evaluation (e.g., Shek & Sun, 2007), qualitative findings based on focus group interviews with program implementers and students (e.g., Shek & Lee, 2008), interim evaluation (e.g., Shek, Sun, & Siu, 2008), analyses of the weekly diaries of students (e.g., Shek, Sun, Lam, Lung, & Lo, 2008), and case studies (e.g., Shek & Sun, 2008). The evaluation findings based on different evaluation strategies indicate that Project P.A.T.H.S. promotes the development of its program participants.

Process evaluation has already been carried out in the EIP and FIP for Secondary 1 (Shek, Ma, Lui, & Lung, 2006) and 2 students (Shek, Lee, & Sun, 2008). The evaluation results indicate that the quality of implementation and program adherence are high. The current study focuses on the Secondary 3 Tier 1 Program.

Process Evaluation for the Secondary 3 Tier 1 Program

Process evaluation for Secondary 3 students is important. First, the Secondary 3 curriculum is different from the others. It requires students to develop self-reflexivity during the process. Thus, the findings of the process evaluation may be different. Second, Secondary 3 students are cognitively more mature and have more life exposure than their Secondary 1 and 2 counterparts. Their perception of program implementation quality can be different. Finally, Secondary 3 students have participated in Project P.A.T.H.S. continuously for 3 years and have all completed the entire Tier 1 curriculum. Their feedback represents an overall evaluation for the entire Project P.A.T.H.S. curriculum.

The current process evaluation focuses on program adherence, process factors, program quality, and success. Program adherence is the objective estimation of the adoption percentage from the manualized plan for real service delivery. A variety of process factors exist. A review of literature indicates that the following program attributes can affect the quality and success of the positive youth development program implementation (Collaborative for Academic, Social and Emotional Learning, 2010; Harachi et al., 1999; Nation et al., 2003; Ringwalt et al., 2003; Tobler, Lessard, Marshall, Ochshorn, & Roona, 1999):

1. *Student interest*: A successful program usually elicits the interest of students.
2. *Active involvement of students*: The more involved the students are, the higher the possibility that the program can achieve its outcomes.
3. *Classroom management*: The program implementer can manage student discipline during student activities. Students obey the requirements set by the program implementer and are attentive.
4. *Interactive delivery method*: Interactive delivery is better than didactic delivery for positive youth development programs.
5. *Strategies to enhance the motivation of students*: The use of various learning strategies can enhance the engagement of students and result in positive learning outcomes.
6. *Positive feedback*: The use of praise and encouragement throughout the lessons by the program implementers can promote the engagement of students.
7. *Familiarity of implementers with the students*: All other things being equal, a higher degree of familiarity with the students is positively related to student learning outcome.
8. *Reflective learning*: The program implementer should engage students in reflection and deeper learning. This can lead to growth and meaningful changes among the students.

9. *Program goal attainment*: The achievement of program goals constitutes program success.
10. *Time management*: Efficient time management ensures that the majority of the program materials are carried out with high program adherence.
11. *Familiarity of program implementers with the implementation materials*: Familiarity with the material ensures that the messages are conveyed effectively to the students.

Program quality is the subjective appraisal of the program implementation process. It can be reflected from the implementation atmosphere and the interaction between program implementers and students.

Program success refers to the extent of unit objective attainment and the subjective evaluation of the response of the students to the program.

Against this background, the current study aims to explore the factors related to the implementation quality and implementation success of the Secondary 3 Tier 1 Program during the Full Implementation Phase. There are two research questions:

1. What is the implementation quality of the Secondary 3 curriculum of the Tier 1 Program of Project P.A.T.H.S. in Hong Kong?
2. How are program adherence and other indicators related to the implementation quality and success of the Secondary 3 Tier 1 Program?

Method

Participants and Procedure

In total, 14 schools were randomly selected from among the 167 secondary schools that joined the Secondary 3 program in the school year 2008/2009 for the process evaluation.

Process evaluation was carried out using systematic observations of actual classroom program delivery. For each school joining the process evaluation, one to two program units were evaluated by two independent observers who are project colleagues with master's degrees. A total of 20 units were observed for this study. The learning units of these units are shown in Table 1. During the observation, observers sat at the back of the classroom and evaluated the method by which the units were actually implemented by completing several instruments.

After the psychometric properties of the instruments were explored, program adherence and implementation process components were associated with implementation quality and success. In addition, program and implementation process components were used to predict implementation quality and implementation success separately.

Instruments

Program adherence. Observers were requested to rate program adherence in terms of percentage (i.e., the correspondence between actual program delivery and stipulated program materials). Pearson correlation analyses showed that the ratings of program adherence were highly reliable ($r = .86$, $p < .001$) between raters.

Implementation Process Checklist (IPC)

The IPC consists of 11 items, which are shown in Table 2. Items 1, 2, 3, 4, 5, 6, and 8 are conceptually related to the implementation process, whereas items 7, 9, 10, and 11 are related to the implementation context (items 7, 9, 10, and 11). Observers were requested to report their observations using a 7-point Likert scale ranging from 1 (*extremely negative*) to 7 (*extremely positive*).

To explore whether the conceptual distinction of these two components is reflective from the data, principal components analysis (PCA) with varimax rotation was used to summarize the effects of the 11 process evaluation items. Two components were identified with eigenvalues greater than 1.0. In addition, the resulting scree plot of the eigenvalues revealed that the leveling off to a straight horizontal line occurred after the second eigenvalue. These two factors could explain 75.02% of variance.

The components emerged to reflect clearly the factors originally proposed. The items from each subscale were loaded on the intended components. Consistent with the conceptual model, two components were formed, namely, implementation process (items 1, 2, 3, 4, 5, 6, and 8) and implementation context (items 7, 9, 10, and 11).

The internal consistency of the overall IPC, as shown by Cronbach's α, was .93. The inter-rater reliability of the IPC, as shown by Pearson correlation, was .87 ($p < .001$). The internal consistency of the Implementation Process subscale was .92 and that of the Implementation Context subscale was .82.

Process Outcomes

Two items were used to evaluate the observation outcome: implementation quality and implementation success. Observers were requested to indicate their observations using a 7-point Likert scale ranging from 1 (*poor*) to 7 (*excellent*). A higher score represents better implementation quality or success. The inter-rater reliability for implementation quality, as shown by Pearson correlation, was .73 ($p < .001$), whereas that for implementation success was .69 ($p < .001$).

Results

The inter-rater reliabilities of the scores were high, allowing the ratings of each item by the two observers across all units to be averaged. Table 3 shows the descriptive profile of the evaluation indicators for process evaluation. The overall program adherence to the established manual ranged from 12.5% to 95.0%, with an average overall adherence of 76.18%. All other items used the 7-point scale. We set 4.50 as the cutoff point as an indication of high or low rating; it is a more stringent criterion instead of using the mid-point. This can differentiate

Table 1. Summary of the Program Objectives of the Observed Units

School	Program Units	Program Objectives
A	MC 3.1	To discuss the differences between fairness in our ideals and in reality
		To understand that a system or situation of "absolute fairness" does not exist in reality
	MC 3.2	To learn how to exercise self-reflection and how to help others
		To discuss ways of helping others in society
B	PN 3.1	To understand that prosocial and moral consideration and analysis are essential when making decisions
C	PN 3.2	To understand that society has different expectations of different roles
		To investigate the potential conflict between being prosocial and socially accepted behaviors
D	RE 3.3	To state how Mencius looked at adversity
		To reflect upon oneself and how Mencius' teachings can be applied in daily life
	RE 3.4	To construct a vision of one's future family
		To recognize that one needs to work hard and use resources properly so as to achieve their aspirations
E	BC3.1 and BC 3.2	To understand the importance of forgiving others sincerely
		To learn how to forgive others for their offenses against us
		To learn how to observe and appreciate people and things around us
F	BF 3.1	To adopt a realistic and positive attitude in exploring future careers
G	BC 3.2	To understand the importance of sincere forgiveness
		To understand the negative influence of taking revenge on those who have offended us
H	BC 3.1	To understand that appreciation brings joy to oneself and others.
		To learn how to observe and appreciate people and things around us, and to express sincere appreciation
		To learn how to respond to appreciation in a proper manner
	BC 3.2	To understand the importance of sincere forgiveness
		To understand the negative influence of taking revenge on those who have offended us
I	BF 3.2	To understand that different jobs have different requirements
		To be aware of the issue of gender stereotypes and their impact(s) on career choices
J	BC 3.1	To understand that appreciation brings joy to oneself and others.
		To learn how to observe and appreciate people and things around us, and to express sincere appreciation
		To learn how to respond to appreciation in a proper manner
	BC 3.2	To understand the importance of sincere forgiveness
		To understand the negative influence of taking revenge on those who have offended us
K	SE 3.1	To understand that successful wealth management relies on the ability to exercise self-control and delayed gratification
		To understand the importance of controlling desires for unnecessary material things
	SE 3.2	To understand the meaning of dreams and their importance in life
		To identify the personal qualities that help one overcome environmental constraints and realize dreams
L	MC 3.3	To learn to cherish love relationships and to love with commitment instead of quitting easily
		To discuss the proper attitudes to end a love relationship
M	SC 3.3	To understand the reasons for conflict among siblings
		To learn the proper attitude to get along with siblings
N	SC 3.3	To understand the reasons for conflict among siblings
		To learn the proper attitude to get along with siblings
	SC 3.4	To understand the reasons for conflict among friends
		To learn how to face and handle conflict with friends

Note. MC = moral competence; PN = prosocial norms; RE = resilience; BC = behavioral competence; BF = beliefs in future; SE = self-efficacy; SC = social competence.

some factors from others and provide a more balanced picture. The scores for implementation quality and success were 4.63 ($SD = .94$) and 4.68 ($SD = .82$), which are high. The scores of the 11 process evaluation items ranged from 4.48 to 5.60. Classroom management (5.60) and familiarity with students (5.40) had the highest scores, whereas reflective learning (4.48) and time management (4.55) had the lowest scores. Apart from reflective learning, all scores were on the high side.

The 11 items were divided into the two groups of the PCA: implementation process and context. The mean score for implementation process was 5.03 ($SD = .97$), whereas that for the implementation context was 4.95 ($SD = .99$). Both scores were on the high side.

Table 4 shows the inter-correlations among program adherence, implementation process, implementation context, implementation quality, and implementation success. All variables were highly related to each other. Quality versus success (.98) and process versus quality (.83) had the highest correlations, whereas process versus adherence (.51) and process versus context (.67) had the lowest.

In addition to correlational analysis, multiple regression analyses were also performed using program adherence, implementation process, and implementation as independent

Table 2. Factor Loadings for Principal Components Analysis With Varimax Rotation of Implementation Process Checklist

Evaluation Items	Implementation Process	Implementation Context
Interest	**.85**	.38
Involve	**.96**	.02
Class	**.71**	.28
Interact	**.74**	.41
Motivation	**.84**	.23
Feedback	**.90**	.22
FStudents	−.02	**.79**
Reflect	**.77**	.45
Goal	.56	**.66**
Time	.32	**.67**
FMaterials	.41	**.82**

Note. Factor loadings > .50 are in boldface.
Interest = student interest; Involve = active involvement of students; Class = classroom management; Interact = interactive delivery method; Motivation = strategies to enhance the motivation of students; Feedback = positive feedback; FStudents = familiarity of implementers with students; Reflect = reflective learning; Goal = program goal attainment; Time = time management; FMaterials = familiarity of program implementers with the program materials.

Table 3. Descriptive Statistics of Evaluation Items

Evaluation Items	Min	Max	*M*	*SD*
Interest	3.0	6.5	5.02	1.02
Involve	3.5	6.5	5.38	.94
Class	4.0	7.0	5.60	.82
Interact	2.5	6.5	4.73	.94
Motivation	3.0	6.5	5.10	1.01
Feedback	2.5	6.5	4.93	1.03
FStudents	3.0	7.0	5.40	1.02
Reflect	2.5	6.0	4.48	1.03
Goal	2.0	6.5	4.83	1.18
Time	2.5	6.5	4.55	.96
FMaterials	2.5	6.5	5.03	.82
Adhere	12.50%	95.00%	76.18%	24.00%
Quality	2.0	6.0	4.63	.94
Success	2.5	6.0	4.68	.82

Note. Interest = student interest; Involve = active involvement of students; Class = classroom management; Interact = interactive delivery method; Motivation = strategies to enhance the motivation of students; Feedback = positive feedback; FStudents = familiarity of implementers with students; Reflect = reflective learning; Goal = program goal attainment; Time = time management; FMaterials = familiarity of program implementers with the program materials; Adhere = program adherence; Quality = implementation quality; Success = implementation success.

variables. Implementation quality and implementation success were used as two separate dependent variables. Table 5 shows the results for the prediction of implementation quality. Both implementation process and program adherence could predict the quality with a large variance explained. Implementation context could not predict the quality. The effect size for the implementation process ($\beta = .51$), Cohen f^2, was .35, which is large. The effect size for program adherence ($\beta = .34$) was .13, which is medium. Table 6 shows the results of the prediction of implementation success. Implementation process, implementation context, and program adherence could all

Table 4. Summary of Intercorrelations for Scores of Process Evaluation, Implementation Quality, and Success

Measure	1	2	3	4	5
1. Implementation process		.67***	.51**	.83***	.81***
2. Implementation context			.77***	.77***	.79***
3. Program adherence				.81***	.78***
4. Implementation quality					.98***
5. Implementation success					

Note. Bonferroni correction was used to evaluate the significance of the correlations.
***$p < .001$,
**$p < .005$.

Table 5. Regression Table of Implementation Quality

Predictors	β
Implementation process	.51***
Implementation context	.22
Program adherence	.34***

Note. $R^2 = .90$.
***$p < .001$.

predict success with a large variance explained. The effect size for implementation process ($\beta = .47$) was .28, which is large. The effect size for both context and adherence ($\beta = .29$) was .10, which is medium.

Discussion and Social Work Implications

Project P.A.T.H.S is a huge evidence-based project of positive youth development in Hong Kong. This article attempts to examine the adherence and quality of implementation of the Tier 1 Program (Secondary 3 curriculum) of Project P.A.T.H.S. in the second year of the FIP.

We find that the range of program adherence is wide from 12.5% to 95.0%. There are various reasons for the difference in the range of program adherence. First, some program materials are overpacked within 30 min. By the time the program implementers have finished part of the program, the lesson time is up. The fact that the implementers cannot complete the program directly affects the adherence rate. Second, the units observed cover a great variety of constructs. Some schools have already some established programs covering these constructs before the P.A.T.H.S. Hence, these schools may use these materials instead of manual materials. For example, the construct of "beliefs about future" is related to career planning (i.e., BF 3.1 and BF 3.2). The research team designed some exercises for students to reflect on their career choices. However, schools may have similar career planning exercises from the career and guidance team and may use their materials instead. Another example is the construct of "moral competence." Unit MC 3.2 is related to prosocial behavior. Almost every school in Hong Kong arranges volunteer services for students (Law & Shek, 2009). Some schools may use their school

Table 6. Regression Table of Implementation Success

Predictors	β
Implementation process	.47***
Implementation context	.29*
Program adherence	.29*

Note. $R^2 = .86$.
*$p < .05$.
*** $p < .001$..

volunteer service experience to deliver similar messages. Thus, the same program objectives are achieved with different materials. Another issue is the use of current examples. The manual was written 2–3 years prior to this study. Thus, some examples may not be up-to-date. Program implementers may opt to choose current cases (such as news clips) rather than the cases provided by the manuals. Although these adaptations aim at a more effective message delivery, they may affect the rate of program adherence.

Despite the discrepancy in the rate of program adherence, the overall degree of adherence to the program units is on the high side. This observation is generally consistent with the previous findings generated from process evaluations conducted by observers (Shek et al., 2006, 2008) and subjective outcome evaluations reported by the program implementers (Shek & Sun, 2008; Shek, Sun, & Siu, 2008). Most program contents are well designed for implementation. This can be attributed to the fact that all program materials have gone through trial teaching. They have already been revised and refined according to prior teaching experience. Thus, program implementers did not have great difficulty in following the plans. These findings dispute the common myth that curricula-based positive youth development programs cannot be used easily and require major adaptations or modifications.

Different aspects of the program delivery were perceived to be positive, highlighting the fact that positive youth programs were well received by both the program implementers and students. Nevertheless, there were relatively low average ratings on time management and reflective learning. These findings are similar to those based on the EIP (Shek et al., 2006, 2008). There are two possible explanations for these observations. First, due to the usual didactic teaching style in Hong Kong, students are not used to reflecting on their everyday life practice in classroom settings. Hence, the students cannot easily shift their learning modes from one-way knowledge dissemination to reflective learning. Second, the overpacking of the curriculum may have prevented the students from carrying out reflections on their learning. Overpacking could have also contributed to the unsatisfactory rating of time management.

The current study has found that program adherence, implementation process, and implementation context are closely associated with implementation quality and success. Implementation quality and success had the highest correlation. For positive youth programs, an interactive program delivery is the key milestone for program quality and success (Collaborative for Academic, Social and Emotional Learning, 2010). Thus, these factors are highly correlated with each other.

Conversely, program adherence is associated with both implementation quality and process. The manualized plan, along with the skills of the program implementer, is effective in bringing quality program delivery. Heavy modification is not required.

Among the correlations, the correlation between implementation process and context was relatively lower. This is consistent with the findings of the PCA that these two variables are distinctive. For example, a stern and distant teacher can achieve all the contextual qualities, such as achieving the goals and familiarity with programs, but this does not reflect good implementation without process-oriented program delivery (Collaborative for Academic, Social and Emotional Learning, 2010).

Program adherence and implementation process had the lowest correlation. This reflects the dilemma of program implementers. In general, a few program materials are overpacked within the time limit. When a program implementer focuses on following the manual through, he or she will run out of time for discussion, self-reflection, and other interactions, affecting the learning process in the end.

The current study utilizes three groups of evaluation variables (i.e., implementation process, implementation context, and program adherence) to predict implementation quality or success. Both process and adherence can predict quality, whereas all variables can predict success. The effect size of all predictors is either medium or large. Implementation process refers to the processes and dynamics during classroom activities, whereas implementation context refers to the background knowledge of students, familiarity with materials, time management, and goal attainment. The implementation process emphasizes the interaction between implementer and students and is critical to the quality and success of positive youth program delivery. Program adherence can predict quality and success, implying that the quality of the curriculum manual is high. Implementation context cannot predict quality, but it can predict success. Context items, including time management, goal attainment, and program preparation, are not closely related to process. However, process is highly related to quality. Thus explains the relatively low relationship to quality. However, implementation success is about whether the message of the lesson is effectively delivered to students and consists of contextual elements. Thus, context can predict implementation success. Other plausible explanations can be explored from further studies.

The findings in the present study have several social work implications. The first implication is on the conceptual level. When we focus on program implementation regardless of external environment (i.e., macro-level implication and dosage), we can focus on three variables: program adherence, implementation process, and context. These variables are all related to implementation quality or success. The present findings provide conceptual insights for understanding program quality or success. These can be generalized to other social services.

For instance, the implementation quality and success of an evidence-based volunteer service manual used by a professional social worker with a group of adolescent volunteers can be evaluated using these three areas.

The second implication is on a practical level. These 11 process evaluation statements can actually be used in other social work contexts, especially in educational and developmental groups. Social workers in general have to develop and implement many group work plans and psychosocial interventions. All these measures are important for positive youth programs and should be brought to the fore in the group work training of social workers working with the youth. Social workers should be made aware that both implementation process and context are important for classroom-based psychosocial intervention programs. With reference to the 11 evaluation items, program implementers should consider ways of enhancing reflective learning of students as well as observe good time management.

Another practice implication is program adherence. O'Connor, Small, and Cooney (2007) suggested that there are certain risky adaptations for program adherence, such as reducing the number or length of sessions, lowering the level of participant engagement, eliminating key messages, removing topics, and changing the theoretical approach. The present findings suggest the importance of program adherence. Program adherence coupled with the effective use of the self, and good interaction between implementers and students, can be more difficult than expected. This requires intensive training and personal reflexivity on the part of the social worker. If overpacking of the material prevents program adherence, then Tier 1 materials can be trimmed down during the revamping process so that the messages can be delivered clearly with sufficient program materials.

This study has several limitations. First, school factors should be controlled in the regression analysis. The authors were not able to locate the academic profile of the schools, as this information is regarded confidential. Thus, the results of regression should be treated with caution and stringently specified in further analysis. Second, only 14 randomly selected schools participated in this study. Although the number of schools can be regarded as respectable, the inclusion of a greater number of schools with different characteristics to participate in the study is advisable. Third, process evaluation with reference to macro-level implication, dosage issues (Saunders, Evans, & Joshi, 2005), and school characteristics can help program developers to understand the quality of the program implementation process further. Fourth, the observation may have a confounding effect. Students may be more cooperative when there are visitors or outside observers because the students do not want to ruin the reputation of their schools. As Chinese students, they also want to "give face" to the program implementers (Leung & Chan, 2003). They intentionally perform better in front of the raters. Fifth, additional variables can be devised for the implementation context and implementation process, such as the effect of using computer games and self-disclosure of the program implementers. Despite these limitations, the current process evaluation findings suggest that the quality of implementation of the Tier 1 Program is generally high. The findings are conducive for program reimplementation as well as the training and conceptual enlightenment of social workers on the importance of process evaluation in social work practice.

Acknowledgments

The preparation for this article and Project P.A.T.H.S. were financially supported by The Hong Kong Jockey Club Charities Trust.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Funding

The authors received financial support from the Hong Kong Jockey Club Charities Trust.

References

Allicock, M., Campbell, M. K., Valle, C. G., Barlow, J. N., Carr, C., & Meier, A., Gizlice, Z. (2010). Evaluating the implementation of peer counseling in a church-based dietary intervention for African Americans. *Patient Education and Counseling*, *81*, 37-42.

Beets, M. W., Flay, B. R., Vuchinich, S., Acock, A. C., Li, K., & Allred, C. (2008). School climate and teachers? Beliefs and attitudes associated with implementation of the positive action program: A diffusion of innovations model. *Prevention Science*, *9*, 264-275.

Bertens, M. G. B. C., Eiling, E. M., van den Borne, B., & Schaalma, H. P. (2009). Uma toril evaluation of an STI/HIV-prevention intervention for Afro-Caribbean women in the Netherlands. *Patient Education and Counseling*, *75*, 77-83.

Bowes, D., Marquis, M., Young, W., Holowaty, P., & Isaac, W. (2009). Process evaluation of a school-based intervention to increase physical activity and reduce bullying. *Health Promotion Practice*, *10*, 394-401.

Braun, S. M., van Haastregt, J. C., Beurskens, A. J., Gielen, A. L., Wade, D. T., & Schols, J. M. (2010). Feasibility of a mental practice intervention in stroke patients in nursing homes: A process evaluation. *BMC Neurology*, *10*, A74.

Carswell, S. B., Hanlon, T. E., O'Grady, K. E., Watts, A. M., & Pothong, P. (2009). A preventive intervention program for urban African American youth attending an alternative education program: Background, implementation, and feasibility. *Education & Treatment of Children*, *32*, 445-469.

Cohen, A. N., Glynn, S. M., Hamilton, A. B., & Young, A. S. (2010). Implementation of a family intervention for individuals with schizophrenia. *Journal of General Internal Medicine*, *25*(Suppl. 1), S32-S37.

Collaborative for Academic, Social and Emotional Learning (2010). *Implementation guidance.* Retrieved February 1, 2011, from http://www.casel.org/implement/guidance.php

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*, 23-45.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation*, *11*, 193-221.

Eisenberg, M. (2009). Integrating a school-based health intervention in times of high-stakes testing: Lessons learned from full court press. *Health Promotion Practice*, *10*, 284-292.

Elliot, D. S., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science*, *5*, 47-52.

Fagan, A. A., Hanson, K., Hawkins, J. D., & Arthur, M. W. (2008). Bridging science to practice: Achieving prevention program implementation fidelity in the community youth development study. *American Journal of Community Psychology*, *41*, 235-249.

Fraze, J. L., Uhrig, J. D., Davis, K. C., Taylor, M. K., Lee, N. R., & Spoeth, S. Robinson, A., Smith, K., Johnston, J., & McElroy, L. (2009). Applying core principles to the design and evaluation of the "Take Charge. Take the Test" campaign: What worked and lessons learned. *Public Health*, *123*(Suppl. 1), e23-e30.

Frazen, S., Morrel-Samuels, S., Reischl, T. M., & Zimmerman, M. A. (2009). Using process evaluation to strengthen intergenerational partnerships in the Youth Empowerment Solutions Program. *Journal of Prevention & Intervention in the Community*, *37*, 289-301.

Gnich, W., Sheehy, C., Amos, A., Bitel, M., & Platt, S. (2008). A Scotland-wide pilot programme of smoking cessation services for young people: Process and outcome evaluation. *British Journal of Addiction*, *103*, 1866-1874.

Griffin, S. F., Wilcox, S., Ory, M. G., Lattimore, D., Leviton, L., & Castro., C. Carpenter, R.A., & Rheaume, C. (2010). Results from the active for life process evaluation: Program delivery fidelity and adaptations. *Health Education Research*, *25*, 325-342.

Harachi, T. W., Abbott, R. D., Catalano, R. F., Haggerty, K. P., & Fleming, C. B. (1999). Opening the black box: Using process evaluation measures to assess implementation and theory building. *American Journal of Community Psychology*, *27*, 711-731.

Hargreaves, J., Hatcher, A., Strange, C., Phela, G., Busza, J., & Kim, J. Watts, C., Morison, L., Porter, J., Pronyk, P., & Bonell, C. (2009). Process evaluation of the intervention with microfinance for AIDS and Gender Equity (IMAGE) in rural South Africa. *Health Education Research*, *25*(1), 27-39.

Hart, A. Jr., Bowen, D. J., Christensen, C. L., Mafune, R., Campbell, M. K., & Saleeba, A. Kuniyuki, A., Shirley A. A., & Beresford, S.A.A. (2009). Process evaluation results from the Eating for a Healthy Life study. *American Journal of Health Promotion*, *23*, 324-327.

Huryk, L. A. (2010). Factors influencing nurses attitudes towards healthcare information technology. *Journal of Nursing Management*, *18*, 606-612.

Johnson, C. C., Lai, Y., Rice, J., Rose, D., & Webber, L. S. (2010). ACTION live: Using process evaluation to describe implementation of a worksite wellness program. *Journal of Occupational and Environmental Medicine*, *52*(Suppl. 1), S14-S21.

Karwalajtys, T., McDonough, B., Hall, H., Guirguis-Younger, M., Chambers, L. W., & Kaczorowsk, J. Lohfeld, L., Hutchison, B. (2009). Development of the volunteer peer educator role in a community Cardiovascular Health Awareness Program (CHAP): A process evaluation in two communities. *Journal of Community Health: The Publication for Health Promotion and Disease Prevention*, *34*, 336-345.

Konle-Parker, D. J., Erien, J. A., & Dubbert, P. M. (2010). Lessons learned from an HIV adherence pilot study in the Deep South. *Patient Education and Counseling*, *78*, 91-96.

Kumpfer, K. L., Pinyuchon, M., de Melo, A. T., & Whiteside, H. O. (2008). Cultural adaptation process for international dissemination of the Strengthening Families Program. *Evaluation & the Health Professions*, *31*, 226-239.

Kwong, K., Ferketich, A. K., Wewers, M. E., Shek, A., Tsang, T., & Tso, A. (2009). Development and evaluation of a physician-led smoking cessation intervention for low-income Chinese Americans. *Journal of Smoking Cessation*, *4*, 92-98.

Law, B. M. F., & Shek, D. T. L. (2009). Adolescent volunteerism in Hong Kong after 1997: A critical review. *International Journal of Adolescent Medicine and Health*, *21*, 137-149.

Leung, T. K. P., & Chan, R. Y. K. (2003). Face, favour and positioning—A Chinese power game. *European Journal of Marketing*, *37*, 1575-1598.

Louis, R. M. S., Parow, J. E., Eby, D. W., Bingham, C. R., Hockanson, H. M., & Greenspan, A. I. (2008). Evaluation of community-based programs to increase booster seat use. *Accident Analysis and Prevention*, *40*, 295-302.

Mair, F. S., Hiscock, J., & Beaton, S. C. (2008). Understanding factors that inhibit or promote the utilization of telecare in chronic lung disease. *Chronic Illness*, *4*, 110-117.

Muckelbuer, R., Libuda, L., Clausen, K., & Kersting, M. (2009). Long-term process evaluation of a school-based programme for overweight prevention. *Child Care, Health & Development*, *35*, 851-857.

Mukoma, W., Flisher, A., Ahmed, N., Jansen, S., Mathews, C., Klepp, K.,, & Schaalma, H. (2009). Process evaluation of a school-based HIV/AIDS intervention in South Africa. *Scandinavian Journal of Public Health*, *37*(Suppl. 2), 37-47.

Nation, M., Crusto, C., Wandersman, A., Kumpfer, K. L., Seybolt, D., Morrissey-Kane, E., & Davino, K. (2003). What works in prevention: Principles of effective prevention programs. *American Psychologist*, *58*, 449-456.

O'Connor, C., Small, S. A., & Cooney, S. M. (2007). Program fidelity and adaptation: Meeting local needs without compromising program effectiveness. *What works, Wisconsin – Research to Practice Series*, *4*, 1-5.

Painter, J. E., Sales, J. M., Pazol, K., Grimes, T., Wingood, G., & DiClemente, R. J. (2010). Development, theoretical framework, and lessons learned from implementation of a school-based influenza vacation intervention. *Health Promotion Practice*, *11*(Suppl. 3), 425-525.

Quintiliani, L., Yang, M., & Sorensen, G. (2010). A process evaluation of tobacco-related outcomes from a telephone and print-delivered intervention for motor freight workers. *Addictive Behaviors*, *35*, 1036-1039.

Ringwalt, C. L., Ennett, S., Johnson, R., Rohrbach, L. A., Simons-Rudolph, A., & Vincus, A. & Thorne, J. (2003). Factors associated with fidelity to substance use prevention curriculum guides in the nation's middle schools. *Health Education & Behavior*, *30*, 375-391.

Salmela, S., Poskiparta, M., Kasila, K., Vahasarja, K., & Vanhala, M. (2009). Transtheoretical model-based dietary interventions in primary care: A review of the evidence in diabetes. *Health Education Research*, *24*, 237-252.

Saul, J., Wandersman, A., Flaspholer, P., Duffy, J., Lubell, K., & Noonan, R. (2008). Research and action for bridging science and practice in prevention. *American Journal of Community Psychology*, *41*, 165-170.

Saunders, R. P., Evans, M. H., & Joshi, P. (2005). Developing a process-evaluation plan for assessing health promotion program implementation: A how-to guide. *Health Promotion Practice*, *6*, 134-147.

Scheirer, M. A. (1994). Designing and using process evaluation. In J. S. Wholey, H. P. Hatry & K. E. Newcomer (Eds.), *Handbook of practical program evaluation.* San Francisco, CA: Jossey-Bass.

Shek, D. T. L. (2006a). Adolescent developmental issues in Hong Kong: Relevance to positive youth development programs in Hong Kong. *International Journal of Adolescent Medicine and Health*, *3*, 341-354.

Shek, D. T. L. (2006b). Construction of a positive youth development program in Hong Kong. *International Journal of Adolescent Medicine and Health*, *18*, 299-302.

Shek, D. T. L., & Lee, T. Y. (2008). Qualitative evaluation of the Project P.A.T.H.S.: Findings based on focus groups with student participants. *International Journal of Adolescent Medicine and Health*, *20*, 449-462.

Shek, D. T. L., Lee, T. Y., & Sun, R. C. F. (2008). Process evaluation of the implementation of the Secondary 2 Program of Project P.A.T.H.S. in the experimental implementation phase. *Scientific World Journal*, *8*, 83-94.

Shek, D. T. L., & Ma, C. M. S. (2011). Impact of the Project P.A.T.H.S. in the junior secondary school years: Individual growth curve analyses. *The Scientific World JOURNAL: TSW Child Health & Human Development*, *11*, 253-266.

Shek, D. T. L., Ma, H. K., Lui, J. H. Y., & Lung, D. W. M. (2006). Process evaluation of the Tier 1 Program of the Project P.A.T.H.S. *Scientific World Journal*, *1*, 300-309.

Shek, D. T. L., & Merrick, J. (2009). Promoting positive development in Chinese adolescents: The Project P.A.T.H.S. in Hong Kong. *International Public Health Journal*, *1*, 237-241.

Shek, D. T. L., & Sun, R. C. F. (2007). Subjective outcome evaluation of the Project P.A.T.H.S.: Qualitative findings based on the experiences of program implementers. *Scientific World Journal*, *7*, 1024-1035.

Shek, D. T. L., & Sun, R. C. F. (2008). Implementation of a positive youth development program in a Chinese context: The role of policy, program, people, process, and place. *Scientific World Journal*, *8*, 960-996.

Shek, D. T. L., Sun, R. C. F., Lam, C. M., Lung, D. W. M., & Lo, S. C. (2008). Evaluation of Project P.A.T.H.S. in Hong Kong: Utilization of student weekly dairy. *Scientific World Journal*, *8*, 13-21.

Shek, D. T. L., Sun, R. C. F., & Siu, A. M. H. (2008). Interim evaluation of the Secondary 2 Program of Project P.A.T.H.S.: Insights based on the experimental implementation phase. *Scientific World Journal*, *8*, 61-72.

Shek, D. T. L., & Yu, L. (2011). Prevention of adolescent problem behavior: Longitudinal impact of the Project P.A.T.H.S. in Hong Kong. *Scientific World Journal, 11,* 546-567.

Shevil, E., & Finlanyson, M. (2009). Process evaluation of a self-management cognitive program for persons with multiple sclerosis. *Patient Education and Counseling*, *76*, 77-83.

Stewart, D. (2008). Implementing mental health promotion in schools: A process evaluation. *The International Journal of Mental Health Promotion*, *10*, 32-41.

Tobler, N. S., Lessard, T., Marshall, D., Ochshorn, P., & Roona, M. (1999). Effectiveness of school-based drug prevention programs for marijuana use. *School Psychology International*, *20*, 105-137.

Wandersman, A., Duffy, J., Flaspholer, P., Noonan, R., Lubell, K., & Stillman, L. Blachman, M., Dunville, R., & Sual, J. (2008). Bridging the gap between prevention research and practice: The interactive systems framework for dissemination and implementation. *American Journal of Community Psychology*, *41*, 171-181.

Wegner, L., Flisher, A., Caldwell, L. L., Vergnani, T., & Smith, E. A. (2008). Health wise South Africa: Cultural adaptation of a school-based risk programme. *Health Education Reserach*, *23*, 1085-1096.

Wilson, D. K., Griffin, S., Saunders, R. P., Kitzman-Ulrich, H., Meyers, D. C., & Mansard, L. (2009). Using process evaluation for program improvement in dose, fidelity and reach: The ACT trial experience. *The International Journal of Behavioral Nutrition and Physical Activity*, *6*, Art79.

Yamada, J., Stevens, B., Sidani, S., Watt-Watson, J., & de Silva, N. (2009). Content validity of a process evaluation checklist to measure intervention implementation fidelity of the EPIC intervention. *Worldviews on Evidence-Based Nursing*, *7*, 158-164.

Zakrzewski, S., Steven, C., & Ricketts, C. (2009). Evaluating computer-based assessment in a risk-based model. *Assessment & Evaluation in Higher Education*, *34*, 439-454.

Zani, B., & Cicognani, E. (2010). Evaluating the participatory process in a community-based health promotion project. *Journal of Prevention & Intervention in the Community*, *38*, 55-69.

Article

American Journal of Evaluation
31(3) 363-381

DOI: 10.1177/1098214010371972
http://aje.sagepub.com

Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation

Brad Astbury[1] and Frans L. Leeuw[2,3]

Abstract

There is growing interest in the concept of "mechanism" across many areas of the social sciences. In the field of program and policy evaluation, a number of scholars have also emphasized the importance of causal mechanisms for explaining how and why programs work. However, there appears to be some ambiguity about the meaning and uses of mechanism-based thinking in both the social science and evaluation literature. In this article we attempt to clarify what is meant by mechanisms in the context of program evaluation by identifying three main characteristics of mechanisms and outlining a possible typology of mechanisms. A number of theoretical and practical implications for evaluators are also discussed, along with some precautions to consider when investigating mechanisms that might plausibly account for program outcomes.

Keywords

theory-driven evaluation, social and behavioral mechanisms, realism, program theory

Introduction

There is now a large and growing body of literature on "mechanisms" and the role of mechanism-based approaches to theory building in the social, behavioral, political, and economic sciences. A key theme in much of this literature is the notion that identifying mechanisms that link cause and effect relations is crucial for the development of deeper and more fine-grained explanations of social phenomena (e.g., Bunge, 1997, 2004; Elster, 1989, 2007; George & Bennet, 2004; Hedström & Swedberg, 1998; Lawson, 1997; Little, 1991; Machamer, Darden, & Craver, 2000; Mayntz, 2004; Steele, 2004; Stinchcombe, 1991; Tilly, 2001).

This widespread interest in mechanism-based explanation across the social science disciplines has slowly started to trickle over into the field of program and policy evaluation. This has occurred

[1]Centre for Program Evaluation, The University of Melbourne, Victoria, Australia
[2]University of Maastricht, Maastricht, The Netherlands
[3]Research, Statistics and Information Center, Ministry of Justice, the Netherlands Institute for Justice, Den Haag, the Netherlands

Corresponding Author:
Brad Astbury, Centre for Program Evaluation, The University of Melbourne, Victoria 3010, Australia
Email: brad.astbury@unimelb.edu.au

mainly through introduction of the term ''mechanism'' into various forms of theory-driven evaluation; most notably Chen's (1990) ''intervening mechanism evaluation'' and the ''realistic evaluation'' of Pawson and Tilley (1997), with its emphasis on middle-range ''context–mechanism–outcome'' theories of programs. Although theory-driven evaluators largely agree that mechanisms are important, there seems to be some lingering confusion about the nature of mechanism-based theorizing. What exactly is a ''mechanism'' and how might the study of ''mechanisms'' be useful for evaluation research?

In this article, we explore these questions by drawing on the extant literature to offer a conceptualization of mechanisms and their role in evaluation research. We begin the article by orientating the reader with a quick history of the now familiar ''black box'' problem in evaluation and describe how this ''problem'' has led to the development of various types of theory-driven approaches to evaluation. Although the focus and form can vary, one key aim of theory-driven evaluation is to unpack programmatic ''black boxes'' and explain how and why programs work (or fail to work) in different contexts and for different program stakeholders. This is where the explicit use of mechanisms can play an important role in assisting theory-oriented evaluators to articulate more precisely the causal linkages between programs and their desired effects.

Next, we discuss at some length the concept of ''mechanism'' and attempt to elucidate what mechanisms are and what they are not. A typology of mechanisms is outlined and illustrated with an example to show how the identification of mechanisms can help evaluators unpack the assumptions underlying an intervention. Finally, we discuss theoretical and practical implications of mechanisms for the field of evaluation, as well as some precautions to consider when using mechanisms to explicate and test program theory.

The Black Box "Problem" and Theory-Driven Evaluation

When evaluators talk about the black box ''problem,'' they are usually referring to the practice of viewing social programs primarily in terms of effects, with little attention paid to how those effects are produced. The antonym of ''black box'' evaluation is white box evaluation (sometimes also referred to as clear box evaluation). White box evaluation involves some attempt to ''unpack'' the black box so that the inner components or logic of a program can be inspected (Scriven, 1994). White box, or ''theory-driven'' evaluation as it is more commonly known, has a long history. Edward Suchman, in his 1967 book *Evaluative Research*, was perhaps the first evaluator to highlight the importance of opening up and empirically testing the ''black box'' of social programs. Weiss (1972, 1995, 1997a), Wholey (1979, 1983), and Chen (1989, 1990), among others, have also helped establish the importance of investigating the theory underlying social programs.

For the theory-driven evaluator, programs are embodiments of theories in at least two ways. First, they comprise an expectation that the introduction of a program or policy intervention will help ameliorate a recurring problem. Second, they involve an assumption or set of assumptions about how and why program activities and resources will bring about change for the better (Tilley, 2004). Another common thread running through the now vast literature on theory-driven evaluation is the concern that programmatic assumptions are often never made explicit and even on occasions when the underlying theory of a program is surfaced, it is not articulated and/or tested in a particularly robust way (Weiss, 1997b). This can be problematic, because if a program is based on a faulty theory, then it will not bring about desired changes, irrespective of how well it is implemented.

Developing Program Theory

In contemporary evaluation practice, program theory is created in many different ways and used for a variety of purposes (Birckmayer & Weiss, 2000; Donaldson, 2007). Program theory can be

developed before a program is implemented (i.e., prospectively) or after the program has been running for some time (i.e., retrospectively). In addition, the way in which program theory is used will usually depend on a variety of factors, such as the particular circumstances of the evaluation; time and resource constraints; the stage of program development; methodological expertise; and the needs of program stakeholders. For example, focusing on program theory could be used as part of an evaluability assessment to help specify the program and determine the feasibility of implementing a full-scale impact study (Wholey, 1987). Alternatively, program theory development may be useful during a participatory evaluation to facilitate team building, staff buy-in, and stakeholder engagement. Policy makers might also benefit from applying theory-driven evaluation to clarify the design of a program prior to implementation and/or establish a performance monitoring framework. In addition, performance auditors can use it to check the robustness of a program's underlying assumptions. Another common application is the development of program theory during evaluation planning to help identify and prioritize key evaluation questions and guide the selection of data collection and analysis techniques.

The methodology for constructing or reconstructing program theory, as well as the level of detail and complexity, also varies significantly. Some examples of these include path analysis and causal modeling; observations of the program in action; interviews with staff to uncover implicit assumptions about how the program works; concept mapping exercises; formal (argumentational) analyses of program and policy documents; and detailed investigations of research on similar programs as well as social science theory (Chen, 1990; Leeuw, 1991; Lipsey, 1993; Trochim, 1989; Smith, 1990). Often evaluators recommend using a combination of these approaches, drawing on both primary and secondary data sources and multiple research methodologies (e.g., Donaldson, 2007, Connell, Kubish, Schorr, & Weiss, 1995; Leeuw, 2003; Pawson & Tilley, 1997).

Program theory is also often expressed in several different ways—a graphic display of boxes and arrows, a table, a narrative description, and so on. There are now many "logical templates" that are widely used by program staff and evaluators to assist in developing visual models of the hypothesized relationship between program resources, activities, and outcomes (e.g., McLaughlin, & Jordan, 1999; United Way of America, 1996; W. K. Kellogg Foundation, 2004). Regardless of the way program theory is developed and depicted, it should constitute a "plausible and sensible model of how a program is supposed to work" (Bickman, 1987, p. 5). It is also argued that theory-driven evaluation should ideally contain both a conceptual component (i.e., the development of the program theory or theories) *as well as* an empirical component (i.e., the systematic testing and refinement of the program theory; Rogers, Hasci, Petrosino, & Huebner, 2000).

Another issue that is commonly raised by proponents of theory-driven evaluation is the need to be more careful in the use of terminology (Davidson, 2006; Weiss, 1997b). Donaldson and Lipsey (2006) note, for example, that the contemporary evaluation landscape is littered with a confusing array of closely related terms such as "theory-based evaluation," "theory-driven evaluation," "program theory," "theory of change," "logic models," "logical frameworks," "intervention logic," and so on.

In particular, while the terms "program theory" and "program logic" are often used interchangeably by evaluators, there appears to be growing recognition that they actually serve different functions (Chen, 2005; Leeuw, 2003; Rogers, 2007; Scheirer, 1987; Weiss, 1997a). Program logic is often used to identify and *describe* the way in which a program fits together, usually in a simple sequence of inputs, activities, outputs, and outcomes. Program theory goes a step further and attempts to build an *explanatory* account of how the program works, with whom, and under what circumstances. Thus, program theory might be seen as an elaborated program logic model, where the emphasis is on causal explanation using the idea of "mechanisms" that are at work.[1]

Although program logic and program theory can indeed be used in such a complementary fashion, this does not seem to be common practice in the field. Arguably, this is because the significance and importance of mechanisms is not well understood by evaluators. As Davidson (2000) has

observed "Despite the purported focus of theory-based evaluation on investigating the causal mechanisms by which a program achieves its effects, surprisingly few actually do this" (p. 18).

Use of the Term "Mechanism" in the Evaluation Literature

Chen and Rossi were among the first evaluators to introduce the term "mechanism" and point out its importance for theory-driven evaluation. For example, in an early article, they argue that "the theory-driven approach avoids the pitfalls of black-box evaluation and provides better understanding of the *causal mechanisms* underlying the relationship between treatment and effects" (Chen & Rossi, 1987, p. 102, emphasis added). Later in Chen's (1990) book, *Theory-Driven Evaluations*, again we see reference to the term "mechanism." This time, a whole chapter is devoted to a particular type of theory-driven evaluation called "intervening mechanism evaluation." This approach involves "identifying the causal processes that theoretically intervene between program treatment and outcome" (p. 191).

More recently, Chen (2005) identifies "two kinds of causal mechanisms [that] may underlie a program: mediating and moderating" (p. 240). He defines these as follows:

> A mediating causal mechanism is a component of a program that intervenes in the relationship between two other components ... [while] the second type of causal mechanism—moderating—represents a relationship between program components that is enabled, or conditioned, by a third factor (pp. 240–241).

Although Chen popularized the term "mechanism" in the early 1990s, in our view, a detailed treatment of the concept of "mechanism" did not appear in the literature until the publication of the book *Realistic Evaluation* by Pawson and Tilley (1997).[2] What makes the approach of Pawson and Tilley to program theory and mechanisms distinctive is that it is strongly based on the principles of causal explanation advanced by early pioneers associated with realist philosophy of science (e.g., Bhaskar, 1975; Harré, 1972). One of the main implications of the realist perspective for evaluation is that it is not enough to simply cite programs as a cause of outcomes—the mechanisms connecting causes and their effects must also be identified. As Pawson and Tilley put it: Programs work (have successful "outcomes") only in so far as they introduce appropriate ideas and opportunities ("mechanisms") to groups in the appropriate social and cultural conditions ("contexts;" p. 57).

A similar point is made by Carol Weiss (1997a) in a popular article reflecting on the past, present, and future of theory-driven evaluation. Weiss argues that it is very important for evaluators to distinguish between what she calls "implementation theory" (or as we would put it "logic models"), which provide operational details about how the program is carried out, and program theory which:

> ... deals with the *mechanisms* that intervene between the delivery of program service and the occurrence of outcomes of interest. It focuses on participants' responses to program service. The mechanism of change is not the program service per se but the response that the activities generate (p. 46, emphasis in original).

So, according to a number of prominent evaluation theorists—Chen, Weiss, Pawson and Tilly and more recently Mark, Henry, and Julnes (2000) and Donaldson (2007)—if evaluators look closely inside the programmatic "black box," they are likely to discover a mechanism or as is more often the case, several mechanism–context–outcome configurations.

Although the concept of mechanism is now part of the official evaluation lexicon, we are less confident that it is understood and applied well in practice. The argument developed in this article is that the potential role for mechanisms in evaluation research, as originally anticipated by several evaluation theorists, appears to have become overshadowed by a misunderstanding of what mechanisms are as well as a narrow focus on linear program logic modeling.[3] In the next section of this article, we look more

closely at why this might be the case and as a possible way forward offer a conceptualization of mechanisms, which is drawn from an extensive reading of the wider philosophical and social science literature.

The Concept of Mechanism

Like so many words that are bandied about, "mechanism" can mean different things depending on the particular field of knowledge and context in which it is used.[4] As Mayntz (2004) observes, " . . . a survey of the relevant empirical and methodological literature soon bogs down in a mire of loose talk and semantic confusion about what 'mechanisms' are" (p. 238). If we are not careful, then, there is a risk that the word 'mechanism' like 'theory' may begin to "obscure rather than create understanding" (Merton, 1968).

In an effort to promote understanding, some scholars have compiled "running lists" of the range of definitions of "mechanism," which can be found in the literature (Gerring, 2007; Hedström, 2005). For example, Mahoney (2001, 2003) identified 24 different definitions of the term mechanism, which he sourced from the writings of 21 different authors.[5] Rather than provide yet another definition, or another list of definitions, we will attempt to elucidate the concept of "mechanism" in two main ways. First, we discuss what mechanisms are not, particularly with respect to how the term is sometimes misunderstood in the context of program evaluation. Second, we identify three key characteristics of mechanisms that are located in various definitions put forward by sociologists, political scientists, evaluators, and philosophers of science in the last 40 years.

A common mistake is for evaluators to conflate the term mechanism with program activity. As noted earlier, we are not alone in pushing for a distinction between underlying "mechanisms of change" and a "to do" list of program activities (Rogers, 2007; Weiss, 1997a). Mechanisms appear too frequently as *unexplained "causal arrows*" that seem to flourish so well in the present climate of enthusiasm with visual logic models. This does not seem to be what theory-driven evaluators had in mind when they introduced the concept of "mechanism" to the evaluation community. It is worth quoting at length Weiss (1997a), who uses the example of a contraceptive counseling program to illustrate the distinction between mechanisms and program activities:

> . . . if counselling is associated with reduction in pregnancy, the cause of change might seem to be the counselling. But the mechanism is not the counselling; that is the program activity, the program process. The mechanism might be the knowledge that participants gain from the counselling. Or it might be that the existence of the counselling program helps to overcome cultural taboos against family planning; it might give women confidence and bolster their assertiveness in sexual relationships; it might trigger a shift in the power relations between men and women. These or any of several other cognitive, affective, social responses could be the mechanisms leading to desired outcomes (p. 46).

Another, more complex concern, is when evaluators *directly* equate mechanisms with variables. Mechanisms are sometimes seen as independent causal variables (i.e., the X in the $X \rightarrow Y$ formula), or more often, treated as an intervening variable or set of mediating or moderating variables that attempt to account for why a statistical correlation exists between an independent and dependent variable (i.e., the Z in the $X \rightarrow Z \rightarrow Y$ formula). This does not distinguish adequately the idea of theoretical mechanisms from a statistical or "variable-centred type of theorising" (Hedström & Swedberg, 1998, p. 15; Mahoney, 2001, 2003; Pawson, 1989). Unlike variables, mechanisms are usually not observable attributes of some unit of analysis. Mechanisms attempt to explain why variables are related. In contrast, a mediator or mediating variable is an attempt to empirically measure the mechanism.[6]

To be sure, there are some parallels between "variable versus theoretical" perspectives on mechanisms. For example, statistical measurement and analysis can help to identify and describe

causal relationships between implementation variables and program outcomes. The results of this kind of quantitative causal modeling then provides the raw material for elaboration of theoretical models of the mechanisms that explain how statistical associations are generated. Although variables and mechanisms can perform complementary functions in evaluation research, it is important to avoid conflating the two. To do so risks losing the explanatory power of mechanisms. Thus, while the ontological status of mechanisms is still contested by philosophers, from a methodological point of view, we prefer to see mechanisms residing at a level of abstraction above variables.

There is considerable support for this view of mechanisms in much of the contemporary literature on mechanism-based analysis. Indeed, regardless of methodological or disciplinary traditions, most scholars who write about mechanisms offer a conceptualization that has been influenced to some extent, by realist accounts of causation (see e.g., George & Bennett, 2004; Henry, Julnes & Mark, 1998; Mahoney, 2003; Pawson & Tilley, 1997). According to this perspective, *mechanisms are underlying entities, processes, or structures which operate in particular contexts to generate outcomes of interest*. There are three essential clues located in a ''realist'' reading of mechanisms. These are that:

1. Mechanisms are usually hidden;
2. Mechanisms are sensitive to variations in context; and
3. Mechanisms generate outcomes.

Mechanisms are Usually Hidden

It is common for realist philosophers and realist evaluation methodologists to emphasize that mechanisms are *underlying* and hence often unobservable or ''hidden'' (Pawson, 2008). This captures nicely the idea that to explain social regularities (or program outcomes), we cannot rely exclusively on repeated observations. Instead, we must go below the ''domain of empirical,'' surface-level descriptions of constant conjunctions and statistical correlations to identify the underlying mechanisms that account for ''regularities in the joint-occurrence of events'' (Bhaskar, 1975, p. 13; Pawson, 1989, p. 157).

A favorite metaphor used to demonstrate the spirit of realist, mechanism-based explanation is the clock. It is not possible to understand how a clock works by examining the surface—the numbers on the face and the movement of the hands. We need to prise the clock open and go beneath the ''surface (observable) appearance'' and delve into the ''inner (hidden) workings'' of the ''balanced spring or the oscillation of caesium atoms'' (Pawson & Tilley, 1997, p. 407).

Although useful in conveying the general tenor of mechanism-based explanation, a limitation of this ''nuts and bolts'' or ''cogs and wheels'' imagery is that it lends itself too easily to criticism on the grounds that human behavior cannot be understood in terms of mechanics or machines (Weber, 2006). Although this may be a reasonable prima facie objection, it does not lead automatically to the conclusion that studying mechanisms is not useful for social researchers and evaluators.

Suppose, for example, the results of successive evaluations reveal that students from disadvantaged backgrounds are statistically less likely than students from more privileged social positions to do well at school. The inquisitive evaluator and astute policy maker might ask: Why is this so? What is it about the nature of disadvantage that leads to underperformance at school? By identifying and testing plausible mechanisms underlying this empirical regularity, we ''move beyond thinking about individual variables and the links between them to considering the bigger picture of action in its entirety'' (Anderson et al., 2006, p. 103).

For instance, we might posit the existence of an underlying mechanism relating to expectations teachers have about disadvantaged students and their academic potential. This mechanism-based

conjecture might then be constructed as a propositional statement in terms of a warrant (Leeuw, 2003) as follows:

> If a teacher expects disadvantaged students to underperform at school, then they will underperform (the proposition). This is because of the principle that expectations, even if initially false, are brought about because of the belief that they are true (the warrant).

To test this mechanism, we might randomly select 20% of students at the start of school term and tell their teachers that these students show unusual potential for intellectual growth. In the classic 1968 study ''Pygmalion in the Classroom'' Rosenthal and Jacobson did just this. They found that 8 months later these ''gifted'' children showed much greater gains in intelligence (as measured by standardized tests) than the remaining children. Rosenthal and Jacobson (1968) considered this to be strong evidence supporting the existence of an educational self-fulfilling prophecy or as others have called it a ''belief-formation'' mechanism (Hedström & Swedberg, 1998).

Basically, a belief-formation mechanism occurs when ''an initially false belief of a situation evokes behaviour that eventually makes the false conception come true'' (Hedström & Swedberg, 1998, p. 18). For instance, if we believe that we are about to meet a disagreeable person, we may approach that person so defensively that we turn them into a disagreeable person (Rosenthal & Jacobson, 1968). Belief-formation mechanisms are well documented and have been used to explain a variety of social phenomena, such as hypnosis; placebo and Hawthorne effects; the failure of banks and stock market crashes; and the processes of racial and religious prejudice (see Merton, 1968; Rosenthal & Jacobson, 1968, pp. 3–30; Schelling, 1998).

Belief-formation mechanisms are, of course, just *one* of potentially several, mechanism-based explanations that might account for why disadvantaged students underperform in school.[7] Belief-formation mechanisms may also interact and meld with other mechanisms. For example, teacher expectancies might shape the nature and quality of teacher–student relations, which in turn influences student motivation toward school (i.e., an academic self-efficacy mechanism). The important point is that mechanisms may not be observable, at least in a direct, empirical sense.

This, however, does not mean that mechanisms are not ''real.'' The vast majority of people do not deny the idea of electricity or gravity because they cannot see it; evidence of effects is usually sufficient (e.g., switching on a light or dropping a rock off a cliff). In general, then, it is possible to make a plausible case for the existence of underlying mechanisms by referring to observable effects *which can only be explained as products of underlying mechanisms*[8] (Sayer, 2000).

Mechanisms are Sensitive to Variations in Context

A second key feature of a realist understanding of mechanisms is that mechanisms are sensitive to variations in context, as well as to the operation of other mechanisms in a particular context. Consider the following logic: In Context A, mechanism (M_1) is not activated. That is, M_1 is dormant; still possessing causal ''tendencies'' or ''capacities'' but not the conditions that ''enable'' it to be triggered. In Contexts B and C, the conditions are conducive to triggering M_1. However, in Context C, no effect or different effects are observed. This could be due to a countervailing mechanism (M_2) that is present in Context C but not in Context B.

The implication of this logic is that mechanisms should not be seen as universal ''covering-laws'' that apply always and everywhere. Even in the physical world, casual ''laws'' vary (e.g., water only boils at 100 C when air pressure is at a certain level; Carter & New, 2004).

This is one of the reasons why mechanisms are often likened (after the sociologist Robert Merton) to middle-range theories that position themselves between universal social laws and description (Pawson, 2000, 2010). For Merton, mechanisms are ''elementary building blocks of middle-range

theories'' (Hedström & Swedberg, 1998, p. 6). Sayer (1981/1998) articulates the contingent nature of mechanisms (i.e., ''causal powers'') as follows:

> Gunpowder has the 'causal power' to explode in virtue of its unstable chemical structure. Copper can conduct electricity because of the presence of free ions in its chemical structure. Whether each of these causal powers are ever 'realised' or 'activated' depends upon contingently related conditions, such as the presence of oxygen, low humidity and a spark in the first case, and an electric current in the second (p. 124).

Recall the earlier example concerning teacher expectations of student performance. Self-fulfilling prophecies (or other belief-formation mechanisms) also do not operate in a deterministic, law-like fashion and may, or may not, be activated depending on the circumstances. Not all teachers, in all places, and at all times will hold the same belief about the educational potential of disadvantaged students. Some teachers, as well as students, may even actively resist the underlying mechanism theory.

Whether the causal ''tendencies'' of a particular mechanism is activated is largely dependent, then, on human reasoning and volition. This is because mechanisms work *through* human agents who have the (cognitive) capacity to think and act in terms of causalities and who also possess other capacities that make things happen. In practical terms, people do not react to programs like billiard balls that are hit; rather ''programs only 'work' if people choose to make them work and are placed in the right conditions to enable them to do so'' (Morén & Blom, 2003; Pawson & Tilley, 1994, p. 294). Thus, a key contextual aspect of the operation of mechanisms in the social world is human interpretation of social structures and events.[9]

This is not to say that mechanisms are necessarily context bound; just that context matters. Lawson's (1997) notion of *demi-regularities* (''demi-regs'') is a useful term to depict the way context can affect mechanisms. For example, it can be argued that in total institutions like prisons ''the strength of weak ties'' mechanism (Granovetter, 1973), which points at the importance of having contacts with others outside personal friendship networks for finding information or jobs, works in a different way compared to what happens outside a prison context. One of the challenges for evaluation research is to detect ''demi-regs,'' which can prevent the belief that every program is unique and idiosyncratic.

Mechanisms Generate Outcomes

A third characteristic of mechanisms is that they generate outcomes. The argument that unobservable causal entities (i.e., mechanisms) produce effects, differs considerably from standard depictions of causation promoted by David Hume and extended by John Stuart Mill. Hume believed that we can only know what we experience. Thus, causation cannot be directly observed—only inferred—by examining patterns of regular contingent relations between events.

This reasoning is familiar to many experimental evaluators who (often deliberately) treat programs as ''black boxes'' and their visible effects as light switches that can be turned on and off. House (1991) explains that one of the unfortunate legacies of Humean causality is that ''things got turned around so that what was real was mistaken to be limited to only what we directly experience ... and anything beyond was discredited as metaphysics'' (p. 4).

In contrast, generative accounts suggest that analysis of causation should not stop with surface events (i.e., the things we can observe through our sense data). This is because reality is stratified, not flat, as Hume implied. This opens up the possibility of understanding the generation of events and regularities (such as program outcomes) *at different layers of reality*.[10] Deep, mechanism-based explanation focuses not only on outcomes themselves, and whether evaluators actually observe them happening, but also the underlying generative mechanisms that produce the

outcomes. Social programs, then, consist not just of what we observe (i.e., program inputs, activities, and outcomes) but also of interactions between mechanisms and contexts, which account for what we observe.

Types of Mechanisms and Levels of Analysis

Although there are not yet repositories on mechanisms, several scholars have provided useful summaries of the research literature on mechanisms. Jon Elster (1989, 2007), for example, draws on insights from neuropsychology to economics and political science to identify and discuss some 20 mechanisms that underlie a range of social phenomena. Similarly, Farnsworth (2007) takes legal arrangements like laws and contracts as a starting point and dissects which types of mechanisms play a role when one wants to understand how these arrangements work. He discusses mechanisms such as the ''slippery slope,'' the ''endowment effect,'' and ''framing effects''.[11] Theoreticians within the social sciences have also contributed to knowledge about mechanisms, as work by Festinger (1950, 1954, 1957), Merton (1968), and Olson (1971) has shown.[12]

In recent years, there have been some preliminary attempts to group mechanisms into common categories. For example, building on James Coleman's (1986, 1990) classic macro–micro–macro model of social action, Hedström and Swedberg (1998) suggest that there are three interrelated types of mechanisms:[13] (a) situational mechanisms; (b) action-formation mechanisms; and (c) transformational mechanisms.

Situational mechanisms operate at the macro-to-micro level. This type of mechanism shows how specific social situations or events shape the beliefs, desires, and opportunities of individual actors. Belief-formation mechanisms such as the self-fulfilling prophecy are a good example of a situational mechanism. This mechanism has been shown to affect the way teachers interact with disadvantaged children in the classroom (as discussed earlier).

Action-formation mechanisms operate at the micro-to-micro level. This type of mechanism looks at how individual choices and actions are influenced by specific combination of desires, beliefs, and opportunities. Leon Festinger's (1957) theory of cognitive dissonance illustrates different types of action-formation mechanisms that are used by individuals to reduce psychological distress that often arises when a person holds two contradictory ideas simultaneously. Smokers, for example, often use techniques of rationalization to avoid quitting despite strong evidence that smoking reduces life expectancy (e.g., ''lung cancer only happens to heavy smokers'').

Transformational mechanisms operate at the micro-to-macro level and show how a number of individuals, through their actions and interactions, generate macro-level outcomes. An example is ''cascading,'' by which people influence one another so much that people ignore their private knowledge and rely instead on the publicly stated judgments of others. The ''bandwagon phenomenon''—the tendency to do (or believe) things because many other people do (or believe) is related to this, as are ''group think,'' the ''common knowledge effect,'' and ''herd behavior'' (Elster, 2007).

Applying the Typology of Mechanisms: A Selected Example

How might this typology be useful for an evaluator who is seeking to open up a program or policy ''black box''? Consider, for example, the case of ''naming and shaming'' interventions as analyzed by Pawson (2006). Over the last decade, policy makers, legislators, the police, and regulators have been increasingly active in implementing a range of naming and shaming programs. For example, dissemination of information via websites and newspapers about organizations that do not comply with rules and regulations (e.g., car manufacturing safety standards). Or establishing bulletin boards and other registration and notification activities about released sex

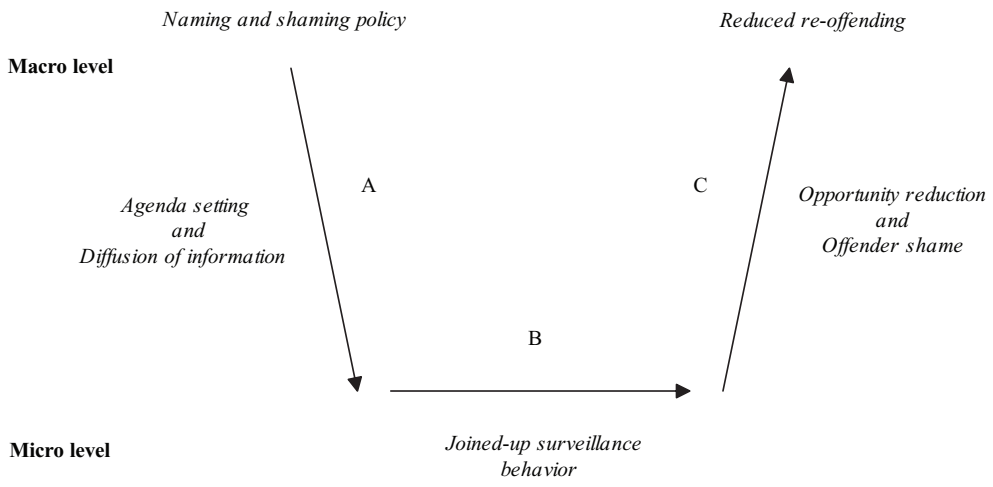


Figure 1. A basic model of mechanisms underlying "naming and shaming" of sex offenders. A = Situational mechanism; B = Action-formation mechanism; C = Transformational mechanism. Source: Adapted from Coleman (1986); Hedström and Swedberg (1998).

offenders, warning the community the ex- offender "is back." It is believed that the naming and shaming of noncompliant and/or deviant behavior will lead to behavior changes in the desired direction.

We focus here on naming and shaming of released pedophilic sex offenders. Examples of interventions in this area are U.S. sex offender registration and community notification initiatives, such as the well-known Megan's Law. This law was enacted in 1996 following the brutal rape and murder of Megan Kanka. Her killer was a released sex offender who, unbeknown to her parents, was living in their neighborhood. News of the case sparked a large public outcry, which resulted in the swift introduction of legislation allowing for a mandatory community notification system for convicted sex offenders (Pawson, 2006). In several other countries such as England, Scotland, the Netherlands, and Australia "look-alike" initiatives are discussed and sometimes implemented.

Applying the typology of mechanisms helps to reveal the (assumed) causal chain of this intervention. The result of our analysis can be found in Figure 1 which presents a basic model of mechanisms underlying naming and shaming policies for pedophilic sex offenders.

There are a range of possible mechanisms at work for each link in the concatenated chain. The links are also likely to be much more complex than the unidirectional arrows in the model suggest. However, for the purpose of illustrating how the typology of mechanisms might assist evaluation practice, we have decided to focus on just a few examples for each level of mechanisms. Where possible, we have also drawn from social science theory to make clear the significance of focusing on different types of putative mechanisms underlying social programs and policies.

The basic model identifies two examples of situational mechanisms, which can affect how sex offender naming and shaming policies work. *Agenda setting* describes the processes by which widespread and extensive media coverage can influence public opinion about how to respond to the issue of released sex offenders living in the community. For example, heightened salience and framing effects are likely to determine the way in which naming and shaming policy is put into practice (McCombs & Shaw, 1972; Scheufele, 1999). In a context of strong social pressure to "do something" about the problem, policy makers and law enforcement officials typically decide that investment in sex offender registration and information management systems are appropriate ways to enact policy.

It is assumed that these activities will then lead to a *diffusion* process, whereby appropriate persons will receive accurate and timely information about registered sex offenders who are residing in their local neighborhood. Diffusion describes how information about registered sex offenders spreads through particular communication channels over time among the members of a social system (Rogers, 2003). In practice, the dissemination of information can occur in a variety of ways. For example, mass media outlets such as the internet as well as interpersonal communication channels such as police-community consultative groups and informal meetings of concerned local citizens.

However, this is not enough to make naming and shaming work. The next link in the chain involves individual actors responding appropriately (or not) to information once it is received. According to the basic intervention theory of naming and shaming, police officers, members of the community, heads of school and teachers, sports mentors, and others will accept information about registered sex offenders, assume that is valid and start to act upon it. Acting upon implies increased *joined-up surveillance behavior*, which also implies that the framing of what suspicious behavior constitutes is more or less similar among different stakeholders in different contexts (Tilley, 1995). It also implies that the results of surveillance activities are fed back into information management systems.

Transformational mechanisms aggregate the individual joint (surveillance) actions to such a level that they are believed to make a difference for the behavioral choices of registered sex offenders. Aggregated surveillance increases the (perceived) likelihood of being caught leading to the possible activation of an *opportunity reduction* mechanism (Clarke, 1992). It also has the potential to trigger a *shame* inducing mechanism among potential re-offenders (Wortley, 1996). Finally, the joining up of the different mechanisms, so the theory goes, will lead to a reduction in re-offending.

This brisk reconstruction of the ''supposed to do'' theory of legislators and policy makers is, of course, just a starting point for planning a potential evaluation of naming and shaming interventions for sex offenders. No doubt there are other possible mechanisms at work that will be discovered as the evaluation process unfolds. Nevertheless, as it currently stands, our preliminary chain of interlinked mechanism-based propositions can be immediately useful in several ways.

First, it places the evaluator in a position to start focusing the evaluation design by identifying questions and data collection methods to ''test'' the way in which the theory works (or fails to work) in practice. For example, situational mechanisms, such as diffusion processes are identified in the policy-maker's theory as an important channel for delivering information to local residents. One way of systematically investigating diffusion would be through a survey approach that considered the extent to which information is actually getting to individuals and community groups in a timely fashion. A local case study analysis of communication networks might also be useful for addressing the question of whether information is actually getting to the right people in the right way. If this is not the case, then unintended side effects such as vigilantism may occur. In both scenarios, data collection efforts are guided by theory, are more focused and address whether the diffusion of information mechanism is working as intended. If not, then the causal chain breaks down and outcomes such as reductions in re-offending will not be achieved.

Second, tapping into existing social science theory relating to mechanisms is often useful for policy makers and evaluator as it may provide important insights into the likely functioning of mechanisms. Consider again the example of situational mechanisms, such as diffusion, where there is a readily available and vigorous body of scholarly research (Rogers, 2003). One important finding is that the type of interpersonal network structure can influence information exchange greatly. Radial personal networks comprise a set of individuals linked to a focal individual but not interacting with each other, while interlocking personal networks consist of individuals who interact mainly with each other. The former would seem to be much more effective in ensuring that information about local sex offenders is exchanged with a wider environment, thereby increasing surveillance

behavior. However, in some communities, this may lead to vigilantism and tighter controls over information networks and communication channels may be required.

A third way in which a more explicit focus on mechanisms in evaluation can be useful is that it helps contribute to knowledge development about social programming more generally. Mechanisms are often "portable" in the sense that they are building blocks for middle-range program theories, which may be transferable to different contexts and policy domains. For example, Pawson (2006) shows that sex offender naming and shaming interventions are not really that unique. The basic idea is also used by policy makers in relation to school league tables and safety indices for car manufacturers, among many others. This suggests that although the precise details and context are likely to vary across different policy situations, this does not necessarily mean that we need to start anew every time a "naming and shaming-type" intervention is launched. Instead of treating all interventions as completely novel, it is our hope that building a knowledge and theoretical base about "families of interventions" (Pawson, 2006), including the different types of mechanisms that underlie them may over time reduce "policy amnesia" and constant reinvention of the wheel.

Using Mechanisms in Evaluation: Some Further Considerations

Although we would argue that evaluators almost always need knowledge about social and behavioral mechanisms to substantiate causal claims, this does not necessarily mean that all evaluations should deal with mechanisms. Certainly, given the potential demand on time and resources, careful consideration must be given to the nature of the evaluand, as well as whether stakeholders actually desire to know how and why the program works. This will provide some clues as to whether an explanatory focus is appropriate. There is nothing inherently wrong with an evaluator deciding to treat a program as a black box, if the purpose of the evaluation is primarily about judging merit, worth, or significance. Sometimes "black box" evaluation is necessary "when no theory [or knowledge of mechanisms] is available or needed, or when getting the relevant expertise would take resources of time or money we do not have" (Scriven, 1998, p. 59). As Scriven might say, even more frankly, "You don't want to waste all your time and resources fishing for mechanisms, when it is not appropriate to be doing so".

If the circumstances are conducive to an explanatory approach, and an evaluator is fortunate enough to have the luxury of "fishing" for mechanisms, then it is important to be aware of some additional precautions. When investigating mechanisms that might plausibly account for program outcomes, it is important not to be "mechanistic." There are no set procedures to be followed rigidly, no columns, rows, or logic model boxes to be filled with generic examples of program "inputs," "outputs," or "activities." The evaluator is not a "box-filling" administrator (Gasper, 2000).

Instead, the evaluator is an applied theorist,[14] who draws on a range of social and behavioral theories to combine diverse ideas in imaginative ways. Weber (2006) summarizes this point well in the context of organizational studies, although it might equally apply to the field of evaluation. He warns that:

> In the quest to identify mechanisms and to assemble them into causal models, researchers end up with theory that is mechanistic and neither interesting nor generative. The temptation is to focus too much on input-output relationships, on linear chains of causality, and on building tightly knit models of arrows and boxes . . . . This would be a dangerous path to take as the complexity and situatedness of much organisational activity begs for a style of theory that preserves some ambiguity (p. 120).

Program theory building with mechanisms involves constant shuttling between theory and empirical data, using both inductive and deductive reasoning. Realist philosophers and evaluators sometimes

refer to this particular logic of inquiry as "retroduction" or "principled discovery" (Bhaskar, 1975; Mark et al. 2000).

It is also important that theory-driven evaluators do not replace substantive social and behavioral science theory with a focus only on putative mechanisms. Again, Weber (2006) is instructive here when he cautions against the temptation to avoid serious theorizing by "plug[ging] and play[ing] with a few handy mechanisms to explain phenomena of interest" (p. 120). Mechanisms "are usually specified in relation to and often only make sense as part of a larger body of theory. They elaborate, sharpen, transpose, and connect theories, but they do not substitute for them" (Weber, 2006, p.120).

Concluding Remarks

Social programs and policies are theories incarnate and a focus on mechanisms in program theory has much to offer the field of evaluation. Theorizing with mechanisms strengthens our understanding of how and why programs work, with whom, and under what circumstances. This is an adage of realist, theory-driven evaluation with which we strongly agree. Although the argument that mechanisms are important for evaluation is not necessarily new, it is one that needs greater emphasis in contemporary evaluation practice.

This is because much of what passes as theory-driven evaluation today seems to pay scant attention to mechanisms (especially in terms of the way we have conceptualized mechanisms in this article). Often, evaluators unpack a programmatic "black box" by laying out the components of the evaluand and then order them into some logical sequence. That is nothing more than dissecting the "operational logic" but not the "conceptual logic," as the U.S. General Accounting Office (US GAO, 1986) once called it. The "operational logic," as we contend, does not constitute an explanatory theory of a program.

This argument has theoretical and practical implications for evaluators. First, a more explicit focus on underlying generative mechanisms might help to counter what appears to be a growing trend toward oversimplified versions of program theory in the form of linear logic models (Gasper, 2000; Rogers, 2007; Weiss, 1997a). The realist notion of explanatory mechanisms is useful for evaluation practitioners who are seeking to complement and extend the way in which they currently develop and use program logic models in evaluation. Evaluators who use primarily quantitative techniques such as path analysis and causal modeling to articulate and test moderators and mediators in program theory might also derive greater explanatory power from their studies if they allow for a distinction between statistical and theoretical conceptions of mechanisms.

Second, while we are optimistic that greater attention to mechanisms will be beneficial for evaluation, it also carries with it some challenges, such as the potential to inadvertently confuse the term "mechanism" with program activity or variable. There is also the danger that mechanism-based explanation becomes associated with a "machine" imagery of social programs or evaluators treat mechanisms as if they are stand-alone little theories that replace the need to engage with substantive social science theory. In this article, we have attempted to address these challenges by reinforcing the argument that mechanism are not a to-do list of program activities nor are they variables that can easily be entered into a statistical regression model. When applied to social policies and programs, mechanisms are the underlying processes or "hidden causal levers" that account for how and why a program works to bring about desired changes in the reasoning and behavior of participants.

Third, mechanism-based theorizing may also help to stimulate and guide much needed research *on* evaluation. This is because evaluation activities themselves also seem to "work" by triggering particular mechanisms in particular contexts. For example, Mark and Henry (2004) identify how evaluation use, or more precisely evaluation influence, is brought about through the interaction of

a range of social and behavioral mechanisms. Mark and Henry argue that their mechanism-based theory of evaluation influence has the potential to open up new areas of inquiry by elucidating how and why evaluation works to bring about desired change in a more precise way. Furthermore, identification and clarification of the mechanisms underling evaluation influence may also provide better guidance for evaluation practitioners who are seeking to improve the influence of evaluation.

In some circumstances, evaluations can also trigger mechanisms that lead to unintended side effects. One example is the mechanism of "designed blindness." It illustrates that stakeholders and evaluators can sometimes become such a strong believer in the program theory they articulate and test that empirical findings are largely, or only, framed as corroborations of this theory (a specific kind of tunnel vision; see Friedman, 2001). Campbell's Law, which relates to the mechanism of "corruptibility of social indicators"[15] (Campbell, 1979), also shows that in some situations, a too strong belief in the relevance of indicators and performance measures and their contribution to "enlightenment" can produce undesired effects such as gaming and goal displacement. Of course, these are just some of the potentially harmful but largely unnoticed mechanisms that lie beneath evaluation activities. Therefore, we anticipate that further research on mechanisms underlying evaluation itself would also be extremely worthwhile.

Finally, getting more involved in (explanatory) theories about mechanisms adds value to the evaluation enterprise because it helps avoid the problem of one-off, discrete evaluations that do little to develop generalizable knowledge about social programming. As policy makers and program officers are not always "au courant" about the mechanisms they assume to be at work, evaluators can demonstrate that although the specific program or intervention they have designed and implemented looks different from other interventions, in fact the same underlying mechanisms are called upon to make the policy work. By sharing and using the accumulated evidence on the level of the *mechanisms at work* (instead of the specific intervention as such), policy makers and evaluator may come to realize that many supposedly "novel" interventions share common underlying mechanisms of change. Knowledge of these mechanisms could then be used to better inform the design and evaluation of social policies and programs.

Notes

1. This useful way of distinguishing between program logic and program theory was suggested to the first author a number of years ago by Dr. Gerald Elsworth.
2. Although the importance of "mechanisms" was discussed at length in an earlier book by Pawson (1989), see also an article by Pawson and Tilley (1994) in the *British Journal of Criminology* where the scientific realist approach to evaluation was foreshadowed.
3. Gasper (2000) has pointed out a number of concerns regarding the overly simplistic application of the "logical framework approach" (or LogFrames) in international development evaluation. He refers, for example, to problems such as "logic-less frames" where prescriptive templates are used leading to an illusion of logic, "lack-frames," which omit critical aspects of a program, and "lock-frames," which restrict program learning and adaptation.
4. For example, Gerring (2007) suggests that in contemporary social science literature there are at least nine distinct, but sometimes contradictory and overlapping, meanings of the term mechanism.
5. Missing from Mahoney's (2003) list of definitions is an important description of mechanisms by the realist philosopher Roy Bhaskar. According to Bhaskar (1975):
   "The world consists of mechanisms not events. Such mechanisms combine to generate the flux of phenomena that constitute the actual states and happenings of the world. They may be said to be real, though it is rarely that they are actually manifest and rarer still that they are empirically identified by men. They are the intransitive objects of scientific theory. They are quite independent of men—as thinkers, causal agents, and perceivers. They are not unknowable, although knowledge of them depends upon a rare blending of

intellectual, practico-technical and perceptual skills. They are not artificial constructs. But neither are they Platonic forms. For they can become manifest to men in experience. Thus, we are not imprisoned in caves, either of our own or of nature's making. We are not doomed to ignorance. But neither are we spontaneously free. This is the arduous task of science: the production of the knowledge of those enduring and continually active mechanisms of nature that produce the phenomena of our world'' (p. 47).

6. This distinction is somewhat similar to the difference between a ''hypothetical construct'' and ''intervening variable'' made by MacCorquodale and Meehl in their seminal 1948 paper.
7. Others mechanisms might include the constraining effects of ''cultural capital'' (Bourdieu & Passeron, 1977) and ''rational choice'' theories of educational inequality (Boudon, 1974).
8. This is, perhaps, what Mahoney (2003) is referring to when he describes mechanisms as ''ultimate causes'' that do not require explanation themselves.
9. Benton (1981/1998) offers a useful distinction between mechanisms in the ''natural world'' and mechanisms in the ''human world.'' In the former, mechanisms are ''person independent,'' whereas in the latter, mechanisms are ''person dependent.''
10. This point relates to questions about the validity of the observation theory that is used when testing a policy theory or an intervention theory. Popper (1972) and Lakatos (1980) have shown that data collection is based on (sometimes) implicit observation theories that can differ in breadth, depth, and width, as is also the case with substantive (i.e., policy) theories. Detecting psychopathic behavior using only a questionnaire is, for example, probably based on an observation theory that is less advanced compared to trying to detect that kind of behavior using magnetic resonance imaging/positron emission tomography (MRI/PET) scans *and* questionnaires. Of course, much more could be said about the relationship between theory and observation, but this would require a lengthy treatment, which is beyond the scope of this article.
11. The ''slippery slope'' mechanism refers to the proposition or argument that a seemingly small first step inevitably leads to a chain of related events culminating in some significant impact. This is analogous to giving a ball a small push on the edge of a downward slope. In economic theory, the ''endowment'' mechanism proposes that people often demand much more to give up something than they would be willing to pay to acquire it. The ''framing effect'' is familiar to opinion researchers who find that presenting the same option in different formats can alter people's decisions. For example, individual inconsistencies have been found depending on whether a question is framed in a negative or positive way.
12. Festinger's cognitive dissonance theory and his social comparison theory (Festinger, 1954, 1957) point at (sociocognitive) mechanisms, whereas Merton's (1973) work on the sociology of science, for example, points at the Matthew effect (explaining how eminent scientists will often get more credit than a comparatively unknown researcher even if their work is similar). Olson's (1971) theory about why large groups do not contribute voluntarily to the production of collective goods is another famous example, as is work by Granovetter (1973) on the strength of weak ties.
13. Essentially, this typology is a categorization of mechanisms according to the way in which they interact at different strata of social reality or levels of analysis. For example, the connection between two macro-level conditions (like the example of low parental income and poor educational outcomes for children) can be explained best by taking into account how macro-phenomena influence the beliefs and actions of individual actors, who over time interact and generate new macro states. As Hedström and Swedberg (1998) explain: This way of conceptualising social action lends itself in a very natural way to a typology of mechanisms . . . instead of analysing relationships between phenomena exclusively on the macro level, one should always try to establish how macro-level events or conditions affect the individual, how the individual assimilates the impact of these macro-level events, and how a number of individuals through their actions and interactions, generate macro-level outcomes (p. 21–22).
14. The notion of an evaluator being an ''applied theorist'' is appealing to us and was first coined by Charles McClintock (1990) in an article describing the use of program theory as part of an evaluation of a local hospice agency.

15. Campbell's law, which was originally proposed in an occasional paper series in 1976, states that "the more any quantitative indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor" (Campbell, 1979, p. 35). Some historical examples of indicators susceptible to these pressures, as cited by Campbell, include "voting statistics," police crime "clearance rates," enemy "body counts" in the Vietnam War, and productivity indicators for factories. A contemporary example would be standardised educational testing in classrooms.

Declaration of Conflicting Interests

The author(s) declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research and/or authorship of this article.

References

Anderson, P. J. J., Blatt, R., Christianson, M. S., Grant, A. M., Marquis, C., Neuman, E. J., & … Sutcliffe, K. M. (2006). Understanding mechanisms in organisational research: Reflections from a collective journey. *Journal of Management Inquiry*, *15*, 102-113.

Benton, T. (1981/1998). Realism and social science: Some comments on Roy Bhaskar's 'The possibility of naturalism.' In M. Archer, R. Bhaskar, A. Collier, T. Lawson, & A. Norrie (Eds.), *Critical realism: Essential readings* (pp. 297-312). London, England: Routledge.

Bickman, L. (1987). The importance of program theory. In L. Bickman (Ed.) *Using program theory in evaluation. New Directions for Program Evaluation* (No. 33, pp. 5–18). San Francisco, CA: Jossey-Bass.

Birckmayer, J. D., & Weiss, C. H. (2000). Theory-based evaluation in practice: What do we learn? *Evaluation Review*, *24*, 407-431.

Bhaskar, R. (1975). *A realist theory of science.* Sussex, Brighton: Harvester Press.

Boudon, R. (1974). *Education, opportunity and social inequality.* New York, NY: John Wiley.

Bourdieu, P., & Passeron, J. C. (1977). *Reproduction in education, society and culture.* London, England: SAGE.

Bunge, M. (1997). Mechanism and explanation. *Philosophy of the Social Sciences*, *27*, 410-465.

Bunge, M. (2004). How does it work? The search for explanatory mechanisms. *Philosophy of the Social Sciences*, *34*, 182-210.

Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*, 67-90.

Carter, B., & New, C. (Eds.). (2004). *Making realism work: Realist social theory and empirical work.* London, England: Routledge.

Chen, H.-T. (1989). The conceptual framework of the theory-driven perspective. *Evaluation and Program Planning*, *12*, 391-396.

Chen, H.-T. (1990). *Theory-driven evaluations.* Newbury Park, CA: SAGE.

Chen, H.-T. (2005). *Practical program evaluation.* Thousand Oaks, CA: SAGE.

Chen, H.-T. & Rossi, P. H. (1987). The theory-driven approach to validity. *Evaluation and Program Planning, 10*, 95-103.

Clarke, R. V. (1992). Introduction. In R. V. Clarke (Ed.), *Situational crime prevention: Successful case studies* (pp. 2-41). Albany, NY: Harrow and Heston.

Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American Journal of Sociology*, *91*, 1309-1335.

Coleman, J. S. (1990). *The foundations of social theory.* Cambridge, MA: The Belknap Press of Harvard University.

Connell, J., Kubish, A., Schorr, L., & Weiss, C. (1995). *New approaches to evaluating community initiatives: Concepts, methods, and contexts.* Washington, DC: Aspen Institute.

Davidson, E. J. (2000). Ascertaining causality in theory-based evaluation. In P. J. Rogers, T. A. Hacsi, A. Petrosino, & T. A. Huebner (Eds.), *Program theory in evaluation: Challenges and opportunities*. *New Directions for Evaluation* (No. 87). San Francisco, CA: Jossey-Bass.

Davidson, E. J. (2006). Editorial: The "baggaging" of theory-based evaluation. *Journal of Multidisciplinary Evaluation*, *3*, iii-xiii.

Donaldson, S. I. (2007). *Program theory-driven evaluation science: Strategies and applications.* Mahwah, NJ: Erlbaum.

Donaldson, S. I., & Lipsey, M. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. In I. Shaw, J. Greene, & M. Mark (Eds.). *The Sage handbook of evaluation* (pp. 56-75). Thousand Oaks, CA: SAGE.

Elster, J. (1989). *Nuts and bolts for the social sciences.* Cambridge, UK: Cambridge University Press.

Elster, J. (2007). *Explaining social behaviour: More nuts and bolts for the social sciences.* Cambridge, UK: Cambridge University Press.

Farnsworth, W. (2007). *The legal analyst: A toolkit for thinking about the law.* Chicago, UK: University of Chicago Press.

Festinger, L. (1950). Informal social communication. *Psychological Review*, *57*, 271-282.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117-140.

Festinger, L. (1957). *A theory of cognitive dissonance.* Evanston, IL: Row, Peterson.

Friedman, V. J. (2001). Designed blindness: An action science perspective on program theory evaluation. *American Journal of Evaluation*, *22*, 161-181.

Gasper, D. (2000). Evaluating the 'logical framework approach:' Towards learning-oriented development evaluation. *Public Administration and Development*, *20*, 17-28.

George, A. L., & Bennett, A. (2004). *Case studies and theory development in the social sciences.* Cambridge, MA: MIT Press.

Gerring, J. (2007). The mechanismic worldview: Thinking inside the box. *British Journal of Political Science*, *38*, 161-179.

Granovetter, M. S. (1973). The strength of weak ties. *The American Journal of Sociology*, *78*, 1360-1380.

Harré, R. (1972). *The philosophies of science: An introductory survey.* London, England: Oxford University Press.

Hedström, P. (2005). *Dissecting the social: On the principles of analytical sociology.* Cambridge, UK: Cambridge University Press.

Hedström, P., & Swedberg, R. (Eds.). (1998). *Social mechanisms: An analytical approach to social theory.* Cambridge, UK: Cambridge University Press.

Henry, G. T., Julnes, G., & Mark, M. M. (Eds.). (1998). *Realist evaluation: An emerging theory in support of practice*. *New Directions for Program Evaluation* (No. 78). San Francisco, CA: Jossey-Bass.

House, E. R. (1991). Realism in research. *Educational Researcher*, *20*, 2-9, 25.

Lakatos, I. (1980). *The methodology of scientific research programmes.* Cambridge, UK: Cambridge University Press.

Lawson, T. (1997). *Economics and reality.* London, England: Routledge.

Leeuw, F. L. (1991). Policy theories, knowledge utilization, and evaluation. *Knowledge and Policy*, *4*, 73-92.

Leeuw, F. L. (2003). Reconstructing program theories: Methods available and problems to be solved. *American Journal of Evaluation*, *24*, 5-20.

Lipsey, M. (1993). Theory as method: Small theories of treatments. In L. B. Sechrest, & A. G. Scott (Eds.), *Understanding causes and generalizing about them. New Directions for Program Evaluation* (No. 57, pp. 5–38): San Francisco, CA: Jossey-Bass.

Little, D. (1991). *Varieties of social explanation: An introduction to the philosophy of social science.* Boulder, CO: Westview Press.

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*, 95-107.

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*, 1-25.

Mahoney, J. (2001). Beyond correlational analysis: Recent innovations in theory and method. *Sociological Forum*, *16*, 575-593.

Mahoney, J. (2003, August 28). *Tentative answers to questions about causal mechanisms*. Paper presented at the annual meeting of the American Political Science Association, Philadelphia, PA. Retrieved October 7, 2008, from http://www.allacademic.com/meta/p62766_index.html

Mark, M. M., & Henry, G. T. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, *10*, 35-57.

Mark, M. M., Henry, G. T., & Julnes, G. (2000). *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs.* San Francisco, CA: Jossey-Bass.

Mayntz, R. (2004). Mechanisms in the analysis of social macro-phenomena. *Philosophy of the Social Sciences*, *34*, 237-259.

McClintock, C. (1990). Evaluators as applied theorists. *Evaluation Practice*, *11*, 1-12.

McCombs, M. E., & Shaw, D. L. (1972). The agenda setting function of mass media. *Public Opinion Quarterly*, *36*, 176-187.

McLaughlin, J. A., & Jordan, G. G. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, *22*, 65-72.

Merton, R. K. (1968). *Social theory and social structure* (3rd ed.). New York, NY: Free Press.

Morén, S., & Blom, B. (2003). Explaining human change: On generative mechanisms in social work practice. *Journal of Critical Realism*, *2*, 37-60.

Olson, M. (1971). *The logic of collective action: Public goods and the theory of groups.* Harvard, UK: Harvard University Press.

Pawson, R. (1989). *A measure for measures: A manifesto for empirical sociology.* London, England: Routledge.

Pawson, R. (2000). Middle-range realism. *Archives Européennes de Sociologie*, *41*, 283-325.

Pawson, R. (2006). *Evidence-based policy: A realist perspective.* London, England: SAGE.

Pawson, R. (2008). Invisible mechanisms. *Evaluation Journal of Australasia, 8*(2), 3-13.

Pawson, R. (2010). Middle range theory and program theory evaluation: From provenance to practice. In J. Vaessen & F. L. Leeuw (Eds.), *Mind the gap: Perspectives on policy evaluation and the social sciences. Comparative Policy Evaluation* (Vol. 16, pp. 171-203). London, England: Transaction Publishers.

Pawson, R., & Tilley, N. (1994). What works in evaluation research? *British Journal of Criminology*, *34*, 291-306.

Pawson, R., & Tilley, N. (1997). *Realistic evaluation.* London, England: SAGE.

Popper, K. R. (1972). *Conjectures and refutations: The growth of scientific knowledge.* London, England: Routledge and Kegan Paul.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.

Rogers, P. J. (2007). Theory-based evaluation: Reflections ten years on. In S. Mathison (Ed.), *New Directions for Program Evaluation* (No. 114, pp. 63–67). San Francisco, CA: Jossey-Bass.

Rogers, P. J., Hacsi, T. A, Petrosino, A., & Huebner, T. A. (Eds.) (2000). *Program theory in evaluation: Challenges and opportunities. New Directions for Evaluation* (No. 87). San Francisco, CA: Jossey-Bass.

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom.* New York, NY: Holt, Rinehart and Winston, Inc.

Sayer, A. (1981/1998). Abstraction: A realist interpretation. In M. Archer, R. Bhaskar, A. Collier, T. Lawson, & A. Norrie (Eds.), *Critical realism: Essential readings* (pp. 120-143). London, England: Routledge.

Sayer, A. (2000). *Realism and social science.* London, England: SAGE.

Scheirer, M. A. (1987). Program theory and implementation theory: Implications for evaluators. In L. Bickman (Ed.), *Using program theory in evaluation. New Directions for Program Evaluation* (No. 33, pp. 59–76). San Francisco, CA: Jossey-Bass.

Schelling, T. C. (1998). Social mechanisms and social dynamics. In P. Hedström & R. Swedberg (Eds.), *Social mechanisms: An analytical approach to social theory* (pp. 32-44). New York, NY: Cambridge University Press.

Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, *49*, 103-122.

Scriven, M. (1994). The fine line between evaluation and explanation. *Evaluation Practice*, *15*, 75-77.

Scriven, M. (1998). Minimalist theory: The least theory that practice requires. *American Journal of Evaluation*, *19*, 57-70.

Smith, N. L. (1990). Using path analysis to develop and evaluate program theory and impact. In L. Bickman (Ed.), *Using program theory in evaluation. New Directions for Program Evaluation* (No. 33, pp. 53–57): San Francisco, CA: Jossey-Bass.

Steele, D. (2004). Social mechanisms and causal inference. *Philosophy of the Social Sciences*, *34*, 55-78.

Stinchcombe, A. L. (1991). The conditions of fruitfulness of theorizing about mechanisms in social science. *Philosophy of the Social Sciences*, *21*, 367-388.

Suchman, E. (1967). *Evaluative research: Principles and practice in public service and social programs.* New York, NY: Russell Sage Foundation.

Tilly, C. (2001). Mechanisms in political processes. *Annual Review of Political Science*, *4*, 21-41.

Tilley, N. (1995). Seeing off the danger: Threat, surveillance and modes of protection. *European Journal on Criminal Policy and Research*, *3*, 27-40.

Tilley, N. (2004). Applying theory-driven evaluation to the British Crime Reduction Programme. *Criminal Justice*, *4*, 255-276.

Trochim, W. M. K. (1989). Outcome pattern matching and program theory. *Evaluation and Program Planning*, *12*, 355-366.

United Way of America. (1996). *Measuring program outcomes: A practical approach.* Alexandria, VA: Author.

U.S. General Accounting Office. (1986). *Teenage pregnancy: 500,000 births a year but few tested programs.* Washington, DC: Author.

Weber, K. (2006). From nuts and bolts to toolkits: Theorizing with mechanisms. *Journal of Management Inquiry*, *15*, 119-123.

Weiss, C. H. (1972). *Evaluation.* Upper Saddle River, NJ: Prentice Hall.

Weiss, C. H. (1995). Nothing as practical as good theory. In J. Connell, A. Kubisch, L. B. Schorr, & C. H. Weiss (Eds.), *New approaches to evaluating community initiatives.* New York, NY: Aspen Institute.

Weiss, C. H. (1997a). Theory-based evaluation: Past, present, and future. *New directions for evaluation.* (No. 76).San Francisco, CA: Jossey-Bass.

Weiss, C. H. (1997b). How can theory-based evaluation make greater headway? *Evaluation Review*, *21*, 501-524.

Wholey, J. S. (1979). *Evaluation: Promise and performance.* Washington, DC: Urban Institute.

Wholey, J. S. (1983). *Evaluation and effective public management.* Boston, MA: Little, Brown.

Wholey, J. S. (1987). Evaluability assessment: Developing program theory. In L. Bickman (Ed.), *Using program theory in evaluation. New Directions for Program Evaluation Theory* (No. 33, pp. 77–92). San Francisco, CA: Jossey-Bass.

W. K. Kellogg Foundation. (2004). *Logic model development guide.* Battle Creek, MI: Author.

Wortley, R. (1996). Guilt, shame and situational crime prevention. In R. Homel (Ed.), *The politics and practice of situational crime prevention* (pp. 115-132). New York, NY: Criminal Justice Press.

INFORMATION SHEET: Foreclosure and Migration into the Rental Housing Market: An Assessment of Adolescents' Wellbeing Following Foreclosure in the District of Columbia The declining housing market and increased unemployment have caused several families to lose their homes to foreclosure. Several hundreds of thousands of families across the United States (US) have lost their homes to foreclosure and as many are at risk of losing their homes to foreclosure. As stressful as losing one's home to foreclosure must be for families, little attention has been devoted to the negative effects foreclosure has on families, particularly adolescents in foreclosed households. This study is an attempt to understand how the stress of foreclosure impacts adolescents' alcohol, marijuana and tobacco use, their depression and their physical stress levels. Participation in this study will involve discussing sensitive issues such as: (1) the experience of being evicted by foreclosure; (2) violence in the neighborhood and in the household; (3) frequency of alcohol, tobacco and marijuana use; and (4) youth's emotional responses to being evicted by foreclosure.

Von Eugene Nebbitt, an Associate Professor in the Jane Addams College of Social Work at the University of Illinois in Chicago, is the person responsible for the study and the safety of participants. The study will be conducted over a **3-year** period. The study is supported by the John D. and Katherine C. MacArthur Foundation.

The primary focus of the study is to understand how the stress of being evicted by foreclosure, adjusting to life in a new neighborhood and perceptions of the new neighborhood influences the physical and mental health, and behavior of adolescents from foreclosed households. However, for the results of this study to be valid we need to compare the results to youth whose families have not been evicted by foreclosure. For this reason, we will recruit three different groups of youth. First, we will recruit youth 13 to 21 years old whose family lost their home to foreclosure in the last year. Second, we will recruit youth 13 to 21 years old in families who live in the same or a similar housing complex as youth whose families were foreclosed. Third, we will recruit youth 13 to 21 years old whose families live in public housing. In addition to being in one of the three groups mentioned above, to be eligible for this study youth must live in Northeast or Southeast, DC, have lived in their current home at least 1 years, do not live with foster parents, and the youth and their consenting parent must be able to read and speak English. If youth are under the age of 18 they will need parent / guardian permission.

Each year of the study youth will be asked to: (1) complete a questionnaire that takes about 45 minutes. A member of the research team will read the questions on the questionnaire and the youth will circle their desired response; and (2) the morning after the youth complete the questionnaire they will ask to give the researchers three small amounts of saliva by using a plastic straw to spat into three small vials. The entire process of depositing the saliva into the three small vials will take approximately one hour and thirty minutes. This activity will take place in the youth's home. Youth will receive a $40 Visa gift card after they complete the questionnaire and we receive the three saliva samples the following morning. Youth will receive the $40Visa Gift Card each year that they complete the questionnaire and give the three saliva samples.

By conducting this study we hope to highlight some of the realities of losing one's home to foreclosure, emphasizing the impact foreclosure has on youth in these households. Ultimately, we hope that the results of this study will be used to create programs and services to make the process of foreclosure less stressful for youth and families. If you have any questions please feel free to contact Dr. Von E. Nebbitt at (314) 550-5735 or the Project Coordinator, Dr. Sharon T. Alston, at (703) 597-2862.

Literature Review Summary Table

Author(s) (year)	Purpose of the study	Research Design	Theory	Research questions/ hypotheses	Sample Characteristics/Size (groups if relevant)	Independent variables (measurement scale)	Dependent variables (measurement scale)	Findings	Limitations of the study

Sampling

For each scenario, identify an appropriate sampling technique and explain why.

1. You are conducting a process evaluation and you want to know how you can improve the delivery of services to your clients. You are trying to improve the CBT program, the mentoring program and the parent enrichment programs.

2. You are interested in knowing what social workers employed at the DC Children and Family Services Agency (CSFA) think about the new legislation "Safe family and Adoption Act". You are interested in knowing how the policy is impacting the clients.

3. You are an evaluator and the funder has asked you do assess if the agency is achieving its outcome. The outcomes are 80% of clients will be employed, 80% of the clients will stop using drugs, 100% will report that that they are satisfied with the program. They want to generalize these findings to the entire population of their clients. They have 5000 clients.

4. You are an evaluator and the funder has asked you do assess if the agency is using the money in the most efficient ways. They also want you to make recommendations about the program. They are interested in knowing if the intake process is working, the discharge process is effective, whether it makes sense to do individual therapy or group therapy. They want you to include clients, but you do not know where all of them are

5. You are piloting a new intervention designed to decrease violence among older people. Violence among older people is a new phenomenon and you don't want to injure the elderly and are unsure of the activities and need to ascertain if the activities are appropriate for the elderly.

5.02 Evaluation and Research

(a) Social workers should monitor and evaluate policies, the implementation of programs, and practice interventions.

(b) Social workers should promote and facilitate evaluation and research to contribute to the development of knowledge.

(c) Social workers should critically examine and keep current with emerging knowledge relevant to social work and fully use evaluation and research evidence in their professional practice.

(d) Social workers engaged in evaluation or research should carefully consider possible consequences and should follow guidelines developed for the protection of evaluation and research participants. Appropriate institutional review boards should be consulted.

(e) Social workers engaged in evaluation or research should obtain voluntary and written informed consent from participants, when appropriate, without any implied or actual deprivation or penalty for refusal to participate; without undue inducement to participate; and with due regard for participants' well-being, privacy, and dignity. Informed consent should include information about the nature, extent, and duration of the participation requested and disclosure of the risks and benefits of participation in the research.

(f) When using electronic technology to facilitate evaluation or research, social workers should ensure that participants provide informed consent for the use of such technology. Social workers should assess whether participants are able to use the technology and, when appropriate, offer reasonable alternatives to participate in the evaluation or research.

(g) When evaluation or research participants are incapable of giving informed consent, social workers should provide an appropriate explanation to the participants, obtain the participants' assent to the extent they are able, and obtain written consent from an appropriate proxy.

(h) Social workers should never design or conduct evaluation or research that does not use consent procedures, such as certain forms of naturalistic observation and archival research, unless rigorous and responsible review of the research has found it to be justified because of its prospective scientific, educational, or applied value and unless equally effective alternative procedures that do not involve waiver of consent are not feasible.

(i) Social workers should inform participants of their right to withdraw from evaluation and research at any time without penalty.

(j) Social workers should take appropriate steps to ensure that participants in evaluation and research have access to appropriate supportive services.

(k) Social workers engaged in evaluation or research should protect participants from unwarranted physical or mental distress, harm, danger, or deprivation.

(l) Social workers engaged in the evaluation of services should discuss collected information only for professional purposes and only with people professionally concerned with this information.

(m) Social workers engaged in evaluation or research should ensure the anonymity or confidentiality of participants and of the data obtained from them. Social workers should inform participants of any limits of confidentiality, the measures that will be taken to ensure confidentiality, and when any records containing research data will be destroyed.

(n) Social workers who report evaluation and research results should protect participants' confidentiality by omitting identifying information unless proper consent has been obtained authorizing disclosure.

(o) Social workers should report evaluation and research findings accurately. They should not fabricate or falsify results and should take steps to correct any errors later found in published data using standard publication methods.

(p) Social workers engaged in evaluation or research should be alert to and avoid conflicts of interest and dual relationships with participants, should inform participants when a real or potential conflict of interest arises, and should take steps to resolve the issue in a manner that makes participants' interests primary.

(q) Social workers should educate themselves, their students, and their colleagues about responsible research practices.